Translated from Russian by Yu. M. Taube
Russian edition edited by Ya. A. Smorodinsky
First published 1987
Revised from the 1983 Russian edition

To the Reader

Mir Publishers would be grateful for your comments on the content, translation and design of this book. We would also be pleased to receive any other suggestions you may wish to make.

Our address is:
Mir Publishers
2 Pervy Rizhsky Pereulok
I-110, GSP, Moscow, 129820
USSR

На английском языке

Printed in the Union of Soviet Socialist Republics

Contents

Introduction

We live as in a mystery dream
On one of convenient planets...

The Russian poet *Igor Severyanin* wrote the above lines in 1909. His idea of a mystery dream implied, I believe, not the lack of scientific explanation of the convenience of our planet. But poet's intuition carried the spirit of his times into his work. This makes readers see more between the lines than the poet himself intended to say. Our understanding of physical phenomena occurring on the Earth has so advanced since, that the state we are in can be no longer called a dream. It is rather the state soon after awakening: logical associations are already established but the consciousness is not yet completely clear.

There is, however, a set of phenomena occurring on the Earth or in the solar system where the cause and effect relationships are quite clear. This book deals with relations and integrity of astronomical and climatic phenomena. These relations are based on celestial mechanics and thermal equilibrium of planets in general and their external containments, atmospheres, for one. The distinguished Soviet geophysicist A. S. Monin introduced the term "geonomy" for the science which deals with a comprehensive analysis of the Earth and its cosmic fellows. Yet this book is no *Geonomy Made Easy* since it does not describe magnetic fields of the planets, as well as electric and optic phenomena in the atmospheres.

We shall make an armchair tour round the Earth and its cosmic surroundings. Our vehicle in this tour will be the knowledge of the basic laws of nature. Theory will guide us to such places which are impossible to be reached even by spacecraft. For example, we shall visit the center of the Sun. However, the description of nature is not our major objective. There are, naturally, some descriptions in this book but they only serve to illustrate the explanations of phenomena. Our wish has been to explain the structure of the world and to demonstrate what makes it so "convenient" to live in.

We begin with pinpointing the exact location of our great home in space and time, namely, describe the part of the universe immediately around us, and the motion of the Earth around the Sun. In Chapter 1 we introduce certain physical notions which are quite complicated compared to those studied at school but which we need for the discussion which follows.

Chapter 2 describes the Earth and planets not as material points but as rotating bodies conforming, however, to forces acting between them. Chapter 3 is focused completely on the Sun, the principal source of energy and the center of attraction in the solar system. Chapter 4 deal with the effect of the solar radiation and chemic composition on the structure of the atmosph and oceans, with the origin of winds and oce currents. Finally, Chapter 5 analyzes the caus of climatic changes on our planet. Amazing climatic alterations over long time intervals depend also on the motion of other planets. So this is the scope of problems discussed in this book.

Chapter 1

The Earth's Path in Space and Time

1. Celestial Sphere

Look at the sky illuminated by countless stars at a clear night. You know that the distance to them is enormous in terms of terrestrial thinking. You already know quite a lot about the arrangement of our world but leave aside the book-knowledge and try to feel the Earth, a tiny grit of the universe rotating and moving in space. Difficult, isn't it? Try again after having read this book.

In this book we treat the Earth in its interaction with the surrounding part of the universe. Therefore we must first of all identify our exact location in space. The celestial sphere, or simply stars in the sky, are our initial frame of reference by which we find direction in space. The variety of stars is very rich. Some of them shine, others are hardly visible. The colour of light may vary from blue to yellow and shades of red. Stars are distributed over the sky quite unevenly: there are some locations where stars are scarce and then comes the Milky Way, a luminous band across the night sky composed of stars which cannot be separately distinguished by a naked eye. The ancient Greeks had a similar name for the Galaxy: Kyklos Galaktos, the milky ring.

If you keep looking at the sky for a long time or memorize the locations of stars and then have

another look at them, you may easily note that the stars have changed positions. This change is however universal: the whole world of stars rotates around us. Of course you know that it is, in fact, the Earth which rotates in relation to stars.

To simplify the orientation in space we should rather disregard the Earth's rotation.

Imagine that you are a cosmonaut observing stars from the orbit of a satellite. It takes you half a turn, less than an hour, to get the view of the entire universe. The dispersed atmospheric light does not interfere and the stars can be seen even near the Sun. Stars do not twinkle in space, the twinkling being the effect of air motion in the Earth's atmosphere. And, what is most important, in space we may disregard the geographical latitude, the time of the day, and the season of the year since the location of stars on the celestial sphere is practically permanent.

Ancient humans observed almost the same view of the celestial sphere: the relative displacement of stars over several thousand years is insignificant, therefore constellations have retained their shapes. The names of many constellations stem from ancient times. Strictly speaking, constellations are 88 arbitrary configurations of the celestial sphere, the boundaries between which were established by the International Astronomical Union between 1922 and 1930. The history of constellations goes back to the groups of stars visible by a naked eye and the ancient names of constellations may stem from the configurations produced imaginarily by linking stars with lines. Drawings of ancient astro-

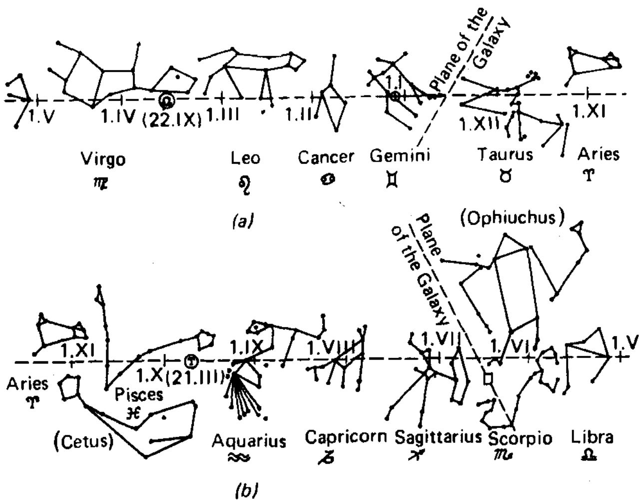


Fig. 1. (*a*) Diagram of Zodiac constellations. The horizontal dashed line is the ecliptic. The point of vernal equinox ♎ and the direction to the Earth's perihelion ⊕ are indicated on the line.
(*b*) Diagram of Zodiac constellations (continued). ♈ is the point of vernal equinox, □ indicates the direction to the center of the Galaxy.

nomers have not survived to our times. Figures 1 and 2 give the shapes of constellations suggested recently by the American astronomer G. Ray. Yet a look at his sketches of Leo and Aquarius makes us think that he has only reproduced ancient drawings by those who had named the constellations. The names have come to us from ancient Greeks, but the Greeks had borrowed the division of the celestial sphere into constellations from the ancient Babylonians. No-

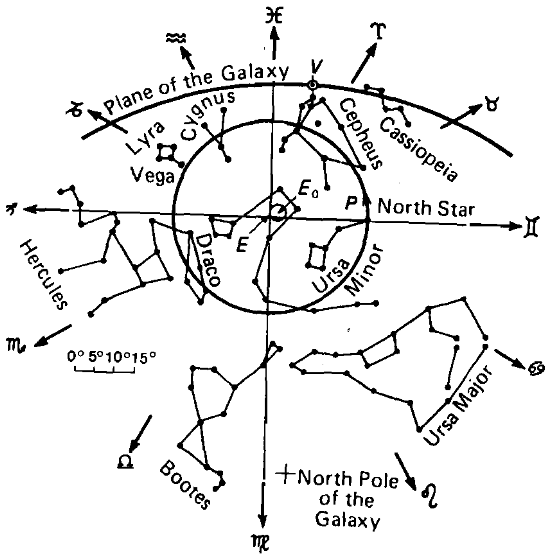


Fig. 2. Diagram of constellations of the northern hemisphere. *P* is the direction to the celestial pole, the larger circle is its trajectory; *E* indicates the pole of ecliptic, the smaller circle is its trajectory; E_0 is the direction of the solar system's momentum; *V* is the direction of the solar system's velocity in the Galaxy; + indicates the north pole of the Galaxy.

tably, in ancient China stars were grouped into constellations absolutely differently.

In Fig. 1 twelve constellations are located along the dotted line to form the Signs of Zodiac: Aries (♈), Taurus (♉), Gemini (♊), Cancer (♋), Leo (♌), Virgo (♍), Libra (♎), Scorpio (♏), Sagittarius (♐), Capricorn (♑), Aquarius

(♒), and Pisces (♓). In brackets there are symbols invented in ancient times to designate these constellations. "Zodiac" is a Greek word meaning "animal" although the Signs of Zodiac include not only animal names.

Figures 1*a* and 1*b* are the developments of a part of the celestial sphere. The dashed line represents the great circle of the celestial sphere. The drawing is closed while in fact the Zodiac constellations are cyclic: Pisces are followed again by Aries. This great circle of the celestial sphere is called ecliptic. The ecliptic is the path of the Sun among the stars during a year as observed from the Earth. Stars cannot be distinguished at day time because of the solar light dispersed in the atmosphere. Yet even the ancient Babylonians understood that stars do not disappear at day and could determine Sun's location in relation to constellations at any given moment of time.

Figure 1 gives not only Zodiac constellations but also two other ones: Ophiuchus and Cetus. Ophiuchus, the Serpent Bearer, is positioned on the ecliptic but it was not included into Zodiac probably on the grounds of its ugly appearance and because of a superstitious wish to make the number of Zodiac constellations "luckier". We have added Cetus because Ray had made it look so attractive and partly also on the grounds of superstition.

For orientation we shall need constellations of the northern hemisphere given in Fig. 2. The center of this figure, point E, is positioned in relation to the ecliptic just as the North pole of the Earth in relation to the equator, while

directions to Zodiac constellations are indicated by arrows. Note that the North Star is not in the center of the figure but near point P.

Astronomers can calculate distances to stars. In fact, stars which compose a constellation are rarely close to one another. In most cases the "neighbouring" stars are separated by enormous distances and their neighbourhood is purely optical.

The same is true for distances from this planet to stars. Thus Vega of Lyra constellation is located at the distance of 2.5×10^{17} m which is 1.5 million times more than that to the Sun. Vega's light reaches our planet in 26.5 years. Such great distances account for the fact that the positions of the stars in constellations are practically constant. In fact, stars travel in relation to one another with measurable velocities. The average velocity of a star is 100 km/s. The figure seems impressive, but let us calculate the time taken by a star moving at such a speed perpendicularly to the direction from us to the star (which is located very close to us, e.g. like Vega) to transit by 1 degree in relation to other, distant, stars. This time is:

$$t = \frac{1^\circ}{180^\circ} \pi \frac{2.5 \times 10^{17}\ \text{m}}{10^5\ \text{m/s}} \simeq 4 \times 10^{10}\ \text{s} \simeq 1400\ \text{years}.$$

There are however stars which transit over the celestial sphere faster than by 1 degree in 1.5 thousand years but such stars are very few. Most stars change their locations much slower since they are situated farther than Vega. For this reason the layout of constellations was practically the same in ancient times as it is today. It

should be noted, however, that the general view may change not only due to the displacement of stars but also because of the alterations in star luminous emittance. It is known, for instance, that Betelgeuse which is presently a red star was mentioned in ancient Chinese chronicles as a yellow star. The general view of a constellation may change radically if only one star turns invisible. This may account for the fact that the ancient names of constellations look strange to the modern eye. Unfortunately, it is yet impossible to calculate the evolution of numerous stars back to ancient times with sufficient accuracy as well as to estimate their luminous emittance at that time.

As mentioned above, stars concentrate on the celestial sphere about the Galaxy. Through a sufficiently powerful telescope one can see that the Galaxy itself is composed of individual stars. However, the projections of these stars on the sphere are so close to one another that a naked eye can distinguish nothing but a continuous luminous cloud. If observed from the Earth, the Galaxy seems to stretch all the way across the sky from horizon to horizon, but for a cosmonaut who can see the entire heaven it is a band of stars around us. The ancient Greeks somehow guessed it: they were the only people to call the Galaxy a circle (Kyklos Galaktos).

Nowadays we use the word "Galaxy" to denote the system of stars housing our Sun and the Earth. We mean thereby not a circle in the celestial sphere but a real three-dimensional formation of stars. We study the Galaxy from the inside but if a sketch of it were made by an outside observer, the resulting picture would have rather an odd shape. The Galaxy looks like a flat and round pancake with a bulge in its center. Spiral "sleeves" extend from the center of the galactic plane and the concentration of stars there is relatively higher. The Galaxy does not have clear boundaries.

The maximum concentration of stars is in the center of the Galaxy, in its nucleus. Unfortunately, the investigation into the nucleus is hindered by interstellar matter which absorbs light. At that place—between Sagittarius and Scorpio—the Milky Way seems to fork over leaving a dark band in the middle. From the center of the Galaxy we receive radio-frequency radiation and short-wave X-radiation; the structure of the galactic nucleus was also studied in infrared light.

Figures 1 and 2 give the galactic plane, the direction to the center of the Galaxy, and the galactic North pole. Sun's place in the Galaxy is near the center of the galactic disk. If it were otherwise the Milky Way would not look like a band around the great circle of the celestial sphere but would seem just a bright spot covering an extensive area. The distance from the Sun to the center of the Galaxy is about $a_{\odot} \simeq 3 \times 10^{20}$ m which by two billion times exceeds that from the Earth to the Sun: $a_{\oplus} = 1.5 \times 10^{11}$ m.

Stars of the Galaxy rotate around its nucleus in conformity with the law of gravitation. The projection of the galactic velocity vector of the Sun is indicated by point V in Fig. 2. The vector lies in the galactic plane, and this means that the Sun has always been in the galactic plane.

The orbital velocity of the solar system $V_\odot \simeq$ 250 km/s. The period of the system's revolution around the galactic center can be assessed as $2\pi a_\odot/v_\odot \simeq 7 \times 10^{15}$ s, i.e. more than two hundred million years.

The law of gravitation permits to calculate the galactic mass inside the Sun's orbit. The total galactic mass approaches this estimation by the order of magnitude:

$$m_g \sim \frac{v_\odot^2 a_\odot}{G} \sim 3 \times 10^{41} \text{ kg},$$

where G is the gravitational constant equal to 6.67×10^{-11} $m^3/(kg \times s^2)$. The total mass of the Galaxy exceeds that of the Sun by a factor of 10^{11}. Approximately this number of stars makes up the Galaxy.

2. The Cosmological Time Scale

The farther from us a cosmic body, the weaker, certainly, its action. However, the weakness of a given action compared to another one is not a sufficient reason to neglect minor disturbances completely. Let us exemplify it.

The greatest force acting upon the Earth is the gravitation of the Sun. Compared to that force, the force of attraction to the center of the Galaxy is insignificant:

$$(a_\oplus/a_\odot)^2 m_g/m_\odot \sim 3 \times 10^{-11}.$$

Is it negligible? The answer depends upon the period of time during which the motion is considered. If it is a period of several years, i.e. several turns of the Earth around the Sun, the effect of galactic gravitation is quite negligible. But if the period in question covers hundreds million years, which is comparable to the period of revolution round the galactic orbit, it is just the weak but constant force of galactic attraction that becomes the major force dictating the Earth's trajectory. The solar gravitation causes only minor alterations of the Earth's trajectory near the galactic orbit of the Sun. Note that the velocity of galactic motion is almost ten times greater than that of the Earth around the Sun.

To study cosmic effects on our planet we may certainly limit ourselves to the Earth's motion in the solar system and the Sun's motion in the Galaxy. But the time scale we shall need for that purpose is larger than the period of revolution round the galactic orbit. Thus we shall focus on the longest time periods, i.e. cosmological time. To do so we shall have first to go back to the description of the surrounding space outside the Galaxy.

There is a great number of equally enormous star systems beside the Galaxy, some of them are similar to ours, others differ significantly. They are called galaxies (with small letter). The nearest two galaxies are Magellanic Clouds observable from the Earth's southern hemisphere. They are removed from us at the distance of 1.6×10^{21} m; the size of each galaxy is about 2×10^{20} m. The Magellanic Clouds are irregular in form and their mass is much less than that of the Galaxy. They are gravitationally connected, i.e. they are Galaxy's satellites.

High-power telescopes permit us to see an enormous number of galaxies, about 10^{11}, removed to

immense distances reaching 10^{26} m. Galaxies are distributed over space quite unevenly, the majority of them being joined into clusters of galaxies. Clusters, in turn, tend to join into superclusters. Nevertheless the universe in general seems to be filled with matter rather evenly; even the number of superclusters in the observable part of the universe is quite significant.

In the early 30s the American astronomer E. Hubble proved by observations that velocities of remote galaxies are directed from us. Moreover, the farther a galaxy, the faster it escapes. The galactic velocities are proportional to the distance to the respective galaxies (this statement is called the Hubble principle). The actual proportionality factor is hard to establish: the universe distances are too long compared to terrestrial measures. The value of the Hubble constant H is about 50-100 km s^{-1}/megaparsec. It demonstrates the rate of recession per megaparsec. The parsec, an astronomical unit of length, eguals the distance at which a base line of one astronomical unit subtends an angle of one second of arc. Thus, it is easy to calculate: megaparsec is equal to $3{,}086 \times 10^{22}$ m. Let us transfer the Hubble constant from astronomical units to the physical with decreased dimensions of length. Then H is approximately 3×10^{-18} s^{-1}.

The idea of the expanding universe implies that all matter of the world was initially packed into a compact superdense agglomeration and then it was hurled in all directions by a cataclysmic explosion. The higher the initial velocity of matter, the farther it has recessed. The fartthest of discovered galaxies move at a speed comparable

with the velocity of light. We shall see further that there is no complete analogy between a conventional explosion and the big-bang expansion of the universe. Yet a natural question may arise: how much time has elapsed since the Big Bang?

To answer this question one must know—beside the Hubble principle—how the universe gravitation decelerates the expansion. This and other problems concerning our world in general are the subject of cosmology. It is quite easy to make an approximate assessment of the age of the universe if gravitation is neglected. Taken the velocities of galactic recession time-independent, we get:

$$t_0 \sim H^{-1} \sim 3 \times 10^{17}\ \text{s} \sim 10^{10}\ \text{years}.$$

More accurate calculations indicate that the age of the universe is between 14 and 20 billion years. The time counted from the start of expansion is called cosmological time.

Amazingly, the extraordinary idea of the expanding world had been theoretically predicted before it was proved by observations. In 1922 the Soviet scientist A. A. Fridman demonstrated that most solutions to Einstein's equations for the world as a whole are unstable and depend on time, and that the expansion of the universe is the most natural effect of gravitational equations. Fridman lived a short life (1888 to 1925) but he carried out a number of most interesting mathematical research and investigations into the theory of the Earth's atmosphere. The final statement of Fridman's work *On the Curvature of Space* (published before the galactic expansion was discovered and the Hubble constant was

first calculated) was the following: "Taken M to be equal to 5×10^{21} masses of the Sun, we get the age of the world to be equal to about 10 billion years." M represents here the mass of the observable universe. Science knows very few examples of such a deep insight!

A clock able of measuring time periods of billion years is also available. For this purpose the radioisotope technique is employed. The technique is based on the instability of isotopes of some chemical elements. Isotopes decay spontaneously and transmute from one to another. The number of radioactive atoms and the mass of the isotope decrease with time in all cases independently of external conditions and in conformity with the following principle:

$$m(t) = m(0)\, 2^{-t/T_{0.5}},$$

where m (0) is the initial mass of the isotope, and $T_{0.5}$ is the half-life, a constant value strictly individual for each isotope. The half-life is the time in which the amount of a radioactive nuclide decays to half its original value.

Half-lives of various isotopes are absolutely different. Short-living atomic nuclei decay in millionth fractions of a second; there are isotopes with $T_{0.5}$ equal to several seconds, the half-life of others may be minutes, days, years. We know presently more than two thousand isotopes of the 107 elements of the periodic table. 305 of them are stable or have half-lives by far exceeding the age of the universe. The distribution of half-lives of other isotopes is given in Fig. 3. It demonstrates that the major part of

unstable isotopes has specific lives between a minute and a week, but there are also some long-livers. The latter are used for radioactive dating.

Nuclear reactions which take place in the cores of stars generate various isotopes of chemical

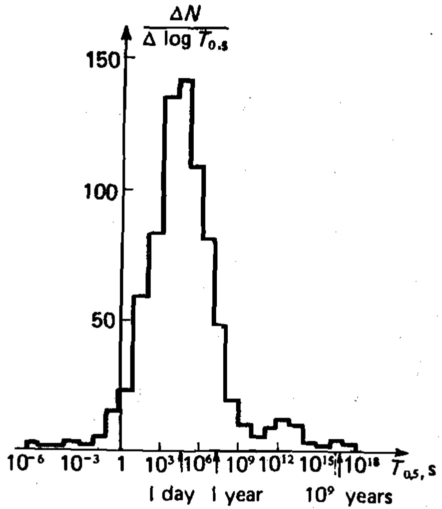


Fig. 3. The distribution of half-lives of isotopes.

elements. We shall discuss some of these processes taking as an example our star, the Sun. There is one more source of unstable isotopes, namely nuclear reactions in the upper layers of the atmosphere induced by fast particles of cosmic rays. That is how Earth's atmosphere, for example, is enriched with carbonic acid which contains carbon isotope ^{14}C. The isotope's half-life is 5570 years. Measuring the content of ^{14}C in wood,

one can calculate the time when that tree was green and growing, when it synthesized organic compounds from the atmospheric carbon dioxide.

Isotopes of star origin with half-lives between 10^5 and 10^7 years have not survived in the Earth's crust. Such isotopes appeared on Earth only after 1942 as a result of controlled nuclear reactions and nuclear explosions.

Finally, several isotopes have half-lives comparable to the age of the universe. These include two uranium isotopes, ^{235}U and ^{238}U, thorium ^{232}Th, potassium ^{40}K, and strontium ^{87}Sr.

A comparison of concentration of these long-livers with the concentration of those they turn into allows to determine the age of oldest rocks, i.e. the time which has passed since the moment of their last melting. For example, out of potassium ^{40}K, which is a component of solid minerals, a gas is slowly formed, argon ^{40}Ar. It is assumed that all argon admixtures have left the sample of rock during the last melting and all argon atoms newly formed from potassium cannot break through the crystal lattice of the solid. Thus the number of argon atoms will be equal to that of the remaining ^{40}K in a rock aged, for instance, 1.26×10^9 years which is the half-life of ^{40}K.

Rocks aged between 10^5 and 10^7 years are dated by way of track counting, i.e. counting the traces of fast particles produced by decay process. The most convenient material for this measurement technique are glassy particles of volcanic tuff and ash. Amazingly, the greatest relative error occurs in dating of rather young rocks aged between 30 and 100 thousand years. These rocks are too young for track techniques, not to mention that volcanic ash falls too rarely over large areas and its concentration is insufficient for an accurate dating by carbon ^{14}C.

The most ancient rocks discovered on Earth are 3.8 billion years old. The age of lunar rocks and meteorites also has a limit: there is no material older than 4.6 billion years in the surrounding part of the solar system. Thus the solar system is believed to have been formed about five billion years ago.

Notably, the age of the solar system determined by the radioisotope technique does not contradict the age of the universe determined by galactic recession. It is slightly less but of the same order of magnitude.

In *The Karamazoff Brothers* by Dostoevsky, a certain Smerdyakov asked Grigory, Karamazoff's servant:

— The Maker created light on the first day and the Sun, the Moon, and stars on the fourth. But where did light shine from on the first day? For which curiosity he was beaten up.

3. Big Bang Light

Had the light shone before the galaxies and stars were formed? Yes. The Big Bang detectable by the galactic expansion had heated the matter of the universe to extremely high temperature. The temperature decreased with expansion, and radiation, which uniformly filled the entire universe, also changed. Yet this primary light still exists. Invisible to the eye, it can be registered by radiotelescopes.

Before we proceed to a deeper discussion of that primary light let us look into the principles of thermal radiation.* You have certainly noted that the more a body is heated the brighter it glows. The chaotic thermal motion of molecules and the frequency of their collisions increase with temperature. The fact is that these phenomena are also accompanied by intensification of chaotic electromagnetic field which we call natural light.

If a body's radiation interacts sufficiently long with the heated medium, it comes to the state of thermal equilibrium. Then the body's properties are determined exclusively by the temperature of its environment. This radiation is called the black-body radiation. Why black? The matter is that thermal equilibrium is achieved if the body absorbs the incident light sufficiently well while the absorbed energy is compensated by the thermal radiation. Bodies which almost completely absorb light of the visible spectrum look black.

Ludwig Boltzmann, a distinguished Austrian physicist of the last century, has established the principle of the thermal radiation: the density of the flux of luminous energy emitted by an absolutely black body is proportional to the fourth power of the temperature:

$$S = \sigma T^4.$$

Flux density S, which is also called radiant intensity, is the energy emitted by a unit of

* You can read about the physics of thermal radiation in more details in the book *Temperature* by Ya. A. Smorodinsky, Mir Publishers, 1984.

body's surface per unit time. Thus, the factor of proportionality σ (the Stefan-Boltzmann constant) is expressed by $J/(m^2 \cdot K^4)$. In 1900 Max Planck, the German physicist, proved the quantum nature of thermal radiation. This allowed us to express the Stefan-Boltzmann constant in terms of fundamental constants: the velocity of light c, the Planck constant $\hbar = 1.054 \times 10^{-34}$ kg·m²/s, and the Boltzmann constant $k = 1.38 \times 10^{-23}$ J/K:

$$\sigma = \frac{\pi^2}{60} \times \frac{k^4}{\hbar^3 c^2} = 5.67 \times 10^{-8}\ \mathrm{kg/(s^3 \cdot K^4)}.$$

Max Planck was occupied with the explanation of the spectrum of thermal radiation. A spectrum is a distribution of luminous intensity over frequencies. It is a function of light frequency ω related to the wavelength $\lambda = 2\pi c/\omega$ demonstrating which part of the energy falls on the frequency range $d\omega$*. Planck was the first to introduce the notion of light quantum, the photon, and employ this new physical idea to explain theoretically the observable spectra of an absolutely black body:

$$\frac{dS}{d\omega} = \frac{\hbar}{4\pi^2 c^2} \times \frac{\omega^3}{e^{\hbar\omega/kT} - 1}.$$

The left-hand side of the equation is the radiant intensity of frequency, ω, related to the spectral interval $d\omega$. Its dimensions are $J/(m^2 \cdot s \cdot s^{-1})$. The seconds can be easily cancelled but the di-

* We could write, of course, a more conventional $\Delta\omega$, the minor frequency range, instead of the differential. The differential should be understood as a small increment. The increment is so small that the variable (ω in this case) and the function do not practically change.

mensions are more characteristic of the value $\frac{dS}{d\omega}$ if written that way. The right-hand denominator includes a power of the number $e = 2.718$, the base of natural logarithms. The thermal spectrum reaches its maximum at a frequency of $\omega_m = 2.82\ kT/\hbar$. If we plot the radiation spectrum $\frac{dS}{d\omega}$ vs frequency, the area under the curve would yield exactly the value of the Boltzmann intensity σT^4. Thus, the intensity of the thermal equilibrium radiation, the frequency of the maximum of its spectrum, and the entire spectral dependence are determined by a sole factor: the temperature.

Figure 4 gives the spectrum of a black-body radiation at 3 K. This is exactly the present temperature of the thermal radiation of the universe. The radiation is a survived evidence of high temperatures at the start of the expansion of the universe. This is why it is called relic, i.e. one that came from the remote past. The relic radiation of the universe was predicted in 1946 by the Russian physicist George Gamow. He estimated the modern temperature of the universe at 10 K which is very close to the true value.

Figure 4 shows the peak value of the spectral curve of the relic radiation to fall on the wavelength of several millimeters. The same electromagnetic radiation occurs in the radio-frequency range undetectable, certainly, by a naked eye. The three-degree black radiation of the universe was discovered by the American astronomers A. Penzias and R. Wilson in 1965.

A natural question may arise: why modern measurements show such a low Big Bang temperature? The fact is that only helium can persist in liquid state at 3 K: the so-called helium temperatures (the extremely low temperatures) have

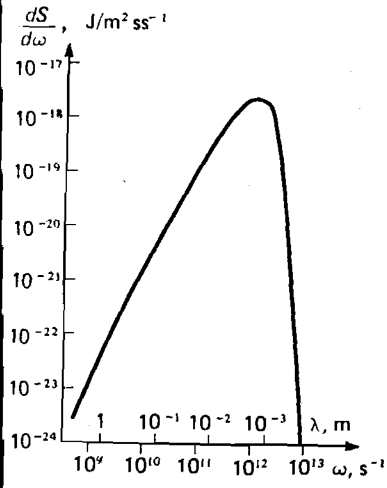


Fig. 4. The spectrum of thermal radiation of the universe with temperature 3 K.

got their name from that specific property of this gas. A temperature so low is a poor match for an idea of explosion. Imagine a radiotelescope receiving the relic radiation. Where does it come to us from? How long ago did it start? Which surface and what material were the source of this radiation?

In this connection we may not fail to mention the uniformity of the universe thermal radiation

in all directions. Whatever part of the heaven the radiotelescope is directed at, it receives radiation of the same temperature varying by mere several thousandths. Yet even these insignificant variations have their own reason and explanation as we shall see below.

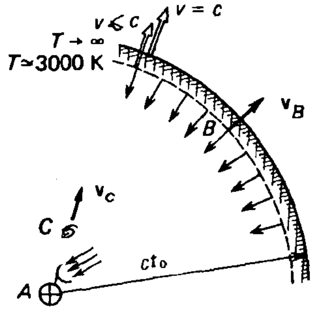


Fig. 5. The diagram of universe expansion and relic radiation propagation.

We shall try to answer the above questions from qualitative positions since their quantitative analysis would require a stricter approach to the notions of cosmology. Consider Fig. 5. The Earth and a radiotelescope receiving the relic radiation are placed arbitrarily at point A. You know already that the modern estimation of cosmological time, t_0, equals approximately 10^{10} years. If we receive today radiation emitted at the moment t_1, we should assume that it has travelled the path $c\,(t_0 - t_1)$. The velocity of light c is the highest possible velocity of information transfer. It is clear that no information can reach us from the distances longer than ct_0.

A part of the sphere of this radius is represented by a continuous arc. The points inside the cross-hatched area are theoretically observable. But the light emitted in this zone is intensively absorbed. This happens because the density and temperature of the material are high at short cosmological time periods: the material is then in the state of plasma and opaque to light.

The recombination of the cosmic plasma, i.e. the composition of electrons and ions into neutral atoms occurred when the universe temperature was approximately 3000 K. The time interval between the *moment of recombination* t_1 and the start of expansion was merely 1 to 1.5 million years. It was then that the universe matter turned transparent from the light-absorbing black. The moment of recombination is expressed in Fig. 5 by the sphere with radius $c\,(t_0 - t_1)$ indicated by dashed lines. It is the radiation of this very surface that is received by radiotelescopes. But why don't we see the red-hot heaven heated to three thousand degrees and do register a temperature thousand times lower?

Recall to mind that the universe does expand. The surface of recombination is relatively close to the limiting sphere with radius ct_0. Thus, it moves away from us at a speed close to the velocity of light. You are familiar with the Doppler effect: if the source of waves moves relatively to the radiation receiver, the received frequency differs from the emitted one. The universe expands, therefore we receive the radiation of expanding galaxies shifted to the red side, the side of longer waves. The surface emitting the relic light moves away very fast at a speed slightly

lower than the velocity of light. For this reason all frequencies of the thermal radiation of this surface decrease by thousand times at 3000 K. In the same way falls the detectable temperature, therefore radiotelescopes "see" the radiation at 3 K.

To understand the composition of the universe one should comprehend the uniformity of all its points: the universe is homogeneous. Figure 5 may lead to an erroneous assumption that we are placed in the center of the world. However, the limiting sphere with the radius ct_0 is not a border of the universe: it is merely an expanding sphere of the information we get about the world. For example, the limiting sphere would be different for an observer placed in galaxy C. To illustrate the homogeneity of the universe we shall describe its evolution once again with the origin of coordinates taken at one of the points from which the relic radiation comes to us at the present moment: at point B. We shall describe in succession a part of the universe at various moments of cosmological time, at the stages of expansion when its structure or composition alter significantly.

We shall start (see Fig. 6a) with the moment only a few seconds after the Big Bang's setoff. In this case the order of magnitude of the temperature is 10^{10} K. At this moment the universe is opaque, the high-temperature radiation emitted from point B being absorbed at a very short distance from the source. Certainly, all points emit radiation, including point A at which our Galaxy would form later. Point A moves away from point B due to expansion. This stage of the

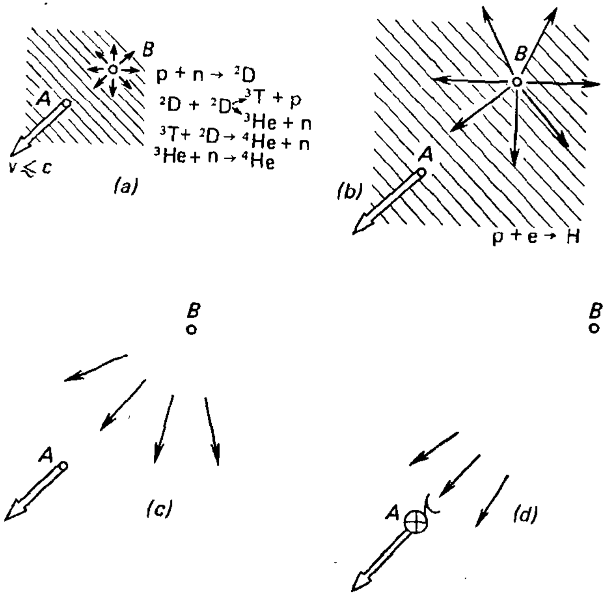


Fig. 6. Propagation of the relic radiation: (a) at the moment of the cosmological synthesis of elements $t \simeq$ 1 s; (b) at the moment of recombination $t \simeq 10^6$ years, $T \simeq 3000$ K; (c) at the stage of the formation of galaxies, $t \simeq 10^9$ years, $T \simeq 15$ to 20 K; (d) modern state: $t \simeq 10^{10}$ years, $T = 3$ K.

evolution of the world is characterized by the formation of the nuclei of the first elements' isotopes (deuterium ^2D, helium ^3He and ^4He, and lithium ^7Li) from protons and neutrons. Theoretical studies indicate that this primary synthesis of elements results in 25 to 27% of isotope ^4He (by mass), and 73 to 75% of the

material remains in the form of protons (hydrogen ^{1}H nuclei). The shares of other resulting isotopes are insignificant: deuterium takes about 10^{-4}, helium ^{3}He accounts for 3×10^{-5}, lithium's share is between 10^{-9} and 10^{-10}. Heavier elements are not formed at that moment: we shall treat the way and time of their coming to this world in Chapter III. Our confidence in the above interpretation of processes which had occurred so long ago is based on the later evidence corroborating just this composition of the universe.

Figure 6*b* gives the universe at the moment of hydrogen's recombination at 3000 K (helium had recombined a bit earlier when the temperature was twice higher). The light is dispersed and absorbed mainly by electrons which are free in neutral atoms. At that moment about half of electrons are still free but the portion of free electrons will decrease rapidly with temperature and the universe would turn transparent.

Figure 6*c* indicates an intermediate moment when the radiation emitted from point *B* has not yet reached point *A*. This point, however, is reached by light from nearer points. From the physical point of view this stage of expansion is marked by the violation of the initial complete homogeneity of the distribution of material caused by gravitation, as a result, clusters of galaxies and galaxies themselves are formed. At that stage of the universe evolution the temperature of the relic radiation was between 10 K and 20 K.

Finally, Figure 6*d* demonstrates the present situation. The radiotelescope receives the relic

radiation from point *B*. As mentioned above, we move from that point at a speed approaching the velocity of light. Thus, in conformity with the Doppler effect we see the spectrum of the universe radiation not at the temperature of 3000 K but of 3 K.

Well, let us arrange now an imaginary experiment. We mount a radiotelescope on a spacecraft which accelerates to a high velocity. Will the relic radiation remain isotropic, i.e. uniform at all sides of the vehicle? No, it will not. The matter is that vehicle's motion creates a motion in a certain direction in the photon gas filling the universe. Therefore the relic radiation will be received with a different Doppler shift depending on the direction of the radiotelescope in relation to the course of the spacecraft. It means that we may measure the spacecraft velocity through the temperature dependence of the relic radiation on the direction.

Stop here. Do we really need a spacecraft? Let us check first whether the Earth itself moves in relation to the black radiation of the universe. Such an experiment was actually carried out in 1979. It has demonstrated that the temperature of the thermal radiation is by 0.1 per cent higher if the radiotelescope is directed at Leo and just as much lower if it faces Aquarius. Hence a conclusion: the solar system moves at a speed of about 400 km/s in relation to the system of coordinates with isotropic temperature of the universe relic radiation. This velocity is called the absolute velocity of the Sun.

We shall treat Earth's orbital motion around the Sun in the next section, and meanwhile we

3-01137

take advantage of some information available on this subject. The matter is that the vector of the absolute velocity of the Sun happens to lie on practically the same plane as the Earth's orbit. Thus the Earth's orbital velocity is added to the absolute velocity of the Sun in winter and is subtracted from it in summer. Consequently, the absolute velocity of the Earth in summer differs from that in winter by 60 km/s and the temperature of the relic radiation in direction to Leo differs from that in direction to Aquarius by 0.54 mK. A comparison of measurements made in December 1980 and July 1981 has revealed that the difference in the temperature of relic radiation approximates the theoretical value. This helped to measure not only the absolute velocity of the Earth but also the yearly alterations of this velocity.

The galactic orbit of the Sun being known, it is possible to calculate the velocity of Galaxy's absolute motion. For this purpose the vector of the orbital velocity of the Sun should be subtracted from the vector of its absolute velocity. So the absolute velocity of the Galaxy approximates 600 km/s. Directions of the above-mentioned motions are given in Fig. 7.

Random deviations of the galactic velocities from the Hubble principle reach approximately the same value: about 600 km/s. Therefore the existence of the absolute velocity of our Galaxy does not contradict the homogeneity and isotropy of the universe which are effective only on the supergalactic scale. Yet the thermal radiation of the universe—taking into account our drift in relation thereto—is isotropic with a high

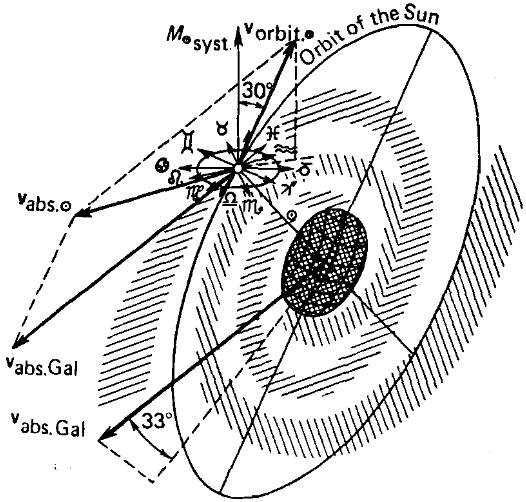


Fig. 7. Diagram of the position of the solar system in the Galaxy and directions of motion of the Sun and Galaxy in relation to the relic radiation.

degree of accuracy; the deviations from isotropy depend upon the accuracy of the experiment.

The calculated absolute velocity of the Earth is of principal importance. Textbooks on physics state that velocity is relative. Should we not doubt it now?

The principle of relativity was introduced by Galileo. It reads that physical principles are identical in all frames of reference moving uniformly and in a straight line. Galileo certainly considered only mechanical experiments when he described in his *Dialogue on Two Systems, the*

Ptolemaean and the Copernican (1632) experiments in a closed cabin inside a ship: "If only the ship's motion is uniform and it does not change course, you would not feel any change in all mentioned phenomena and by none of them you would be able to judge whether the ship is moving or standing still".

However, the Galilean principle remained also true for electromagnetic phenomena. The Michelson's experiment proved in 1881 that the velocity of light is independent of the frame of reference thereby having discarded Lorentz's theory of a motionless ether. The ether was regarded as a certain hypothetical medium in which electromagnetic modes propagate. Electrodynamic laws and the Galilean principle were combined by Albert Einstein into the theory of relativity. The notion of ether lost then its meaning. In memory of the past scientific search and errors the American astrophysicist P. J. Peebles has called the displacement relative to the thermal radiation of the universe "the new ether drift". The experiment made in 1979 was also his idea.

Note however that the measurement of the absolute velocity of the Earth does not refute the Galilean principle since measurements are made not in a closed cabin but in the entire observable part of the universe. The measurement of the absolute velocity also does not contradict the principle of relativity because the velocity in relation to the relic radiation is insufficient to introduce the absolute reference frame of spatial coordinates. Let us make sure of that.

If the absolute velocity of the Galaxy $v_{\text{abs.Gal}}$ is known, we can theoretically indicate the point in the modern universe from which our Galaxy had escaped, i.e. the point where the matter of the Galaxy was at the moment of the Big Bang. The direction to it lies somewhere in Pegasus (it concerns certainly only direction and not the stars of the constellation). The distance to that point is approximately $t_0 v_{\text{abs.Gal}} \simeq 2 \times 10^{24}$ m. There is no sense, however, in the assumption that this point was the ground zero of the Big Bang.

Imagine a distant galaxy, for example, at point C in Fig. 5. It is moving away from us at a speed much higher than 600 km/s but at the same time it may be motionless in relation to the relic radiation. We could as well believe that this galaxy is the plane the world is expanding from. Therefore there is no natural reference point in the universe.

The effect of the relic radiation on our planet is extremely weak: its source is far away and the temperature is low. This effect could have been neglected but the description of the Earth's motion would be incomplete without considering the absolute velocity. Let us return now from remote areas of the universe to the immediate surroundings of our planet—the solar system.

4. The Motion of the Earth and the Planets Around the Sun

The orbital motion in the Galaxy—not to mention the new ether drift—is insignificant for our further discussion. Thus, it is convenient to take the

center of mass of the solar system as the reference point. Since 99.87 per cent of system's mass is concentrated in the Sun itself, the reference point would naturally coincide with the center of the Sun.

Nine major planets (Mercury, Venus, the Earth, Mars, Jupiter, Saturn, Uranus, Neptune, Pluto) and a multitude of minor celestial bodies rotate around the Sun. Notably, the orbital planes of all planets except Pluto almost coincide. The plane of the Earth's orbit is certainly the first priority of our discussion. As you already know, the ecliptic is the projection of the Earth's orbit on the celestial sphere. The majority of observable events which take place in the solar system occurs at the ecliptic* "against the background" of Zodiac. Note that the plane of Moon's rotation around the Earth slopes towards the ecliptic by the angle less than 6°. The Moon is always projected on Zodiac constellations and is seen between them. Therefore the ecliptic is the most convenient reference point of the celestial sphere for the study of the solar system.

The first five planets besides the Earth have been known to the humanity since ancient times. Their slow and mysterious motion among Zodiac constellations has always attracted the eyes of men. These planets, the Sun and the Moon, were believed in ancient times to have magic and mystical properties. They were associated with the seven days of the week, metals known to the ancients at that time, and mythical gods. Ancient

* The word "ecliptic" means "the line of eclipses" in Greek.

Chinese idea was that planets correspond not only to the days of the week and metals but also to five elements. The same idea was popular in ancient Rome where each planet had its own day of the week and metal which fact is presently reflected in many European languages. All this curiosities are given in Table 1. The table carries no physical meaning but it has a certain historical and educational import. Celestial bodies are grouped in Table 1 in the ascending order of their periods of revolution according to Zodiac with corresponding days of the week arranged in an alternate sequence.

Table 1

Planets and Associated Notions

Celestial body	Astronomical symbol	Day of the week	Metal	Element
Moon	☾	Monday	silver	—
Mercury	☿	Wednesday	mercury	water
Venus	♀	Friday	copper	metal
Sun	☉	Sunday	gold	—
Mars	♂	Tuesday	iron	fire
Jupiter	♃	Thursday	tin	wood
Saturn	♄	Saturday	lead	earth

Twelve months of the year correspond to twelve Zodiac constellations. The word "month" is likely to stem from lunar phases (new moon, first quarter, full moon, last quarter). The ancient calendars were timed by new moons. There are 12.37 lunar months in a year. The number was later rounded off to 12 and this could have been the reason for grouping stars into constellations. Note that the division of Zodiac into twelve constellations is in no way natural for the celestial sphere. On the average the angular distances between stars of the adjacent Zodiac constellations are the same as those within the constellations. In ancient China, by the way, there were distinguished 28 Zodiac constellations, the reason therefore could well have been that the same number of days was counted in a lunar month.

Ancient humans did not know that the Earth's motion was the actual cause of the Sun's travel through Zodiac constellations during a year. Thus, a tradition originated in ancient Babylon relates months of the year to Zodiac constellations on which the Sun was projected when observed from the Earth. This relation, however, is not fully identical with the modern one: since then the dates of the Sun's entering the successive Zodiac constellations have shifted significantly. We shall compare the ancient and present data and discuss the shift somewhat later. Yet it should be mentioned that the shift results from the Earth's own rotation.

To trace the real way of the Earth's transit during a year it would be much more convenient to consider not the Sun itself but the direction of the vector coming from the reference point,

i.e. from the Sun to the Earth. In December and January this radius vector is directed at Gemini, in January and February it enters Cancer, and so forth. The points corresponding to these directions on the first days of each month are given in Fig. 1.

Projections of the motion of other planets on the celestial sphere look equally simple in the frame of reference with the Sun as the reference point: all the planets travel quite uniformly in a straight line from left to right (see Fig. 1). In relation to an observer placed on the Earth their trajectories on the celestial sphere are highly complicated: from time to time the forward motion is succeeded by the reverse one. Copernicus was the first to have guessed the reason for that: the intricacy of the planetary motion over the celestial sphere results from the combination of the motion of planets and the Earth's motion around the Sun.

So far we have been treating the projection of the motion of planets on the celestial sphere. Ancient astronomers focused their efforts mainly on the construction of kinematic models of these motions. The next historical step was to measure the distances between the planets and the Sun, to describe planetary orbits.

The great German scientist Johannes Kepler (1571-1630) has established that planets revolve by closed curves called ellipses. For further discussion we shall need some mathematical properties of the ellipse.

If a circular cylinder were cut obliquely or a circle were considered at an angle, the resulting figure would form an ellipse. The ellipse is de-

scribed by the following equation:

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1,$$

where a and b are two parameters of the ellipse called semiaxes (Fig. 8). If $a = b$ the equation of the ellipse turns into the equation of the circle with radius a.

It is however more convenient to characterize an ellipse by its major semi-axis a and a dimen-

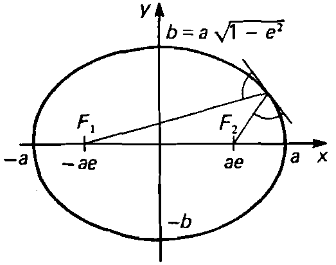


Fig. 8. Ellipse.

sionless number $e = \sqrt{1-\frac{b^2}{a^2}}$ which is called the eccentricity of the ellipse. Evidently, $0 \leqslant e \leqslant 1$. If $e = 0$ is the case, it corresponds to the circle. The greater the eccentricity, the more extended the ellipse. When $e = 1$ it degenerates into a segment with length $2a$.

There are two remarkable points inside the ellipse, the focuses. They are located symmetrically on the major axis, the distance between the focuses is $2ae$. A ray emitted from one focus reflects from the ellipse and comes through the other focus. There is one more feature: the sum

of the distances from any point of the ellipse to its focuses does not depend on the location of the point and is equal to $2a$. These properties of the ellipse focuses are easily proved with the help of the equation of the ellipse.

Compute the area of the ellipse using a method which is not very strict from mathematical point

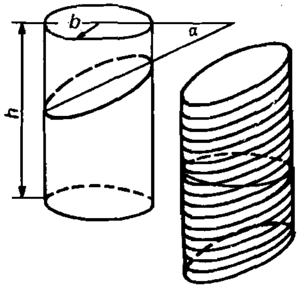


Fig. 9. Deriving area of an ellipse.

of view but rather elegant. Imagine a usual circular cylinder with radius b and height h. You know that its volume is equal to $\pi b^2 h$. Cut now this cylinder by a plane at an angle α. The resultant cross-section will be an ellipse with semi-axes b and $a = b/\cos\alpha$ (Fig. 9).

Attach the cut-off top of the cylinder to its bottom from below. It is clear that the volume of the compound cylinder remains the same. To calculate the volume of this cylinder, we employ the Cavalieri principle: cut the cylinder mentally into a great number n of elliptical disks parallel to new bases. The volume of each of the disks is its area (the yet unknown area of the

ellipse) multiplied by the disk's height $h \cos \alpha/n$. Thus,

$$\pi b^2 h = nS \frac{h \cos \alpha}{n}.$$

Hence, by cancelling out h and expressing $\cos \alpha$ as a, we have the area of the ellipse $S = \pi ab$.

This method of calculating the area of an ellipse was introduced in the 17th century by the Japanese scientist Tacacatsu Seki who had independently discovered the Cavalieri principle. Unfortunately, there was no contact between European and oriental sciences at that time.

Having analyzed long-term observations of the planetary motion, Kepler derived three laws which describe the kinematics of the motion of the solar system bodies:

1. Every planet moves in an ellipse with the Sun at a focus.
2. The radius vector drawn from the Sun to the planet sweeps out equal areas in equal times (i.e. the areal velocity is constant).
3. The squares of the times taken to describe their orbits by two planets are proportional to the cubes of the major semi-axes of the orbits, i.e.

$$\frac{T^2}{a^3} = \text{const}.$$

Several decades later Isaac Newton proved that Kepler laws stem unambiguously from dynamics of the material body motion if: (1) the celestial bodies gravitate to one another; (2) gravitation is directed along the straight line connecting the centers of masses; (3) this force is inversely

proportional to the square of the distance between the bodies.

The connection of these three statements with the Kepler laws is not obvious. We can demonstrate the third Kepler law for circular orbits. Unfortunately, the proof of this law for elliptic orbits and Kepler's first law is extremely complicated. Thus it has to be taken for granted.

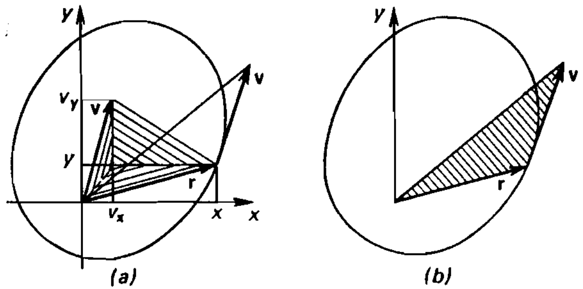


Fig. 10. Deriving the constancy of areal velocity.

Yet there is a way to demonstrate the second Kepler law with the help of the second statement.

The areal velocity, as we may see, equals half of the area of the parallelogram constructed on vectors **v** and **r**. We designate it by A (it is shown by hatching in Fig. 10). Then we project radius vector **r** and velocity **v** onto the coordinate axes; the projections will be x, y, and v_x, v_y. The area of triangle A is calculated as the sum of areas of three smaller triangles hatched differently:

$$A = \frac{(x - v_x)\, y}{2} + \frac{(v_y - y)\, v_x}{2} + \frac{(v_y - y)(x - v_x)}{2}$$
$$= \frac{x v_y - y v_x}{2}.$$

Consider now a change in the value of A resulting from a slight increment in time dt. The definition of velocity implies that the increments in coordinates are: $dx = v_x dt$ and $dy = v_y dt$. Therefore

$$dA = \frac{1}{2}(v_x v_y\, dt + x\, dv_x - v_y v_x\, dt - y\, dv_x)$$
$$= \frac{1}{2}(x\, dv_y - y\, dv_x).$$

In conformity with Newton's second law the change in velocity dv is proportional to the force. The force is directed towards the center, towards the reference point. Thus the ratio between increments in velocities dv_x/dv_y equals the ratio x/y. Hence the increment dA is identically equal to zero, A is constant which fact proves the second Kepler law.

Let us find the areal velocity of the Earth. Parameters of the Earth's orbit around the Sun are well known (see Table 2). During a year the Earth describes an ellipse whose area is $S = \pi ab$. Thus the areal velocity of the Earth can be found by dividing this area by the period of revolution ($\oplus$ is the astronomical symbol of the Earth):

$$A = \frac{S}{T_\oplus} = \frac{\pi a_\oplus^2 \sqrt{1-e^2}}{T_\oplus} = 4.46 \times 10^{15}\ \text{m}^2/\text{s}.$$

The orbital velocity increases as the planet approaches the Sun. The point of the orbit which is the nearest to the Sun is called perihelion (from the Greek words *peri* meaning "near" and "helios", the "sun"). For the Earth's satellites this term is substituted by perigee (from the Greek word *gē*, the "earth"). The distance from

Table 2

Orbits of Planets

Planet	a (in $a_\oplus$)	Period (years)	Mass (in $m_\oplus$)	Moment (in $M_\oplus$)	Inclination of orbit	e
Mercury	0.39	0.24	5.6×10^{-2}	3.4×10^{-2}	7°0′14″	0.2056
Venus	0.72	0.62	8.1×10^{-1}	7.0×10^{-1}	3°23′39″	0.0068
Mars	1.52	1.88	1.1×10^{-1}	1.3×10^{-1}	1°51′0″	0.0934
Jupiter	5.20	11.87	3.2×10^{2}	7.6×10^{2}	1°18′21″	0.0484
Saturn	9.54	29.46	9.5×10^{1}	2.9×10^{2}	2°29′25″	0.0557
Uranus	19.18	84.01	1.5×10^{1}	6.4×10^{1}	0°46′23″	0.0472
Neptune	30.06	164.8	1.7×10^{1}	9.5×10^{1}	1°46′28″	0.0086
Pluto	39.44	247.6	2.0×10^{-3}	1.2×10^{-2}	17°8′38″	0.2486

$a_\oplus = 1.4959787 \times 10^{11}$ m; $T_\oplus = 3.1558150 \times 10^{7}$ s; $m_\oplus = 5.976 \times 10^{24}$ kg; $M_\oplus = 2.66 \times 10^{40}$ kg m^2/s; $e_\oplus = 0.0167$

the Sun to the perihelion is $a_\oplus(1 - e)$. The velocity at perihelion is perpendicular to the radius vector, thus it can be easily found:

$$v_{\max} = \frac{2\pi a_\oplus}{T_\oplus}\sqrt{\frac{1+e}{1-e}} = 30.3\ \text{km/s}.$$

The farthest from the Sun point of the orbit is called aphelion, the distance to it is $a_\oplus(1 + e)$. The constancy of the areal velocity suggests that the velocity at aphelion is

$$v_{\min} = \frac{2\pi a_\oplus}{T_\oplus}\sqrt{\frac{1-e}{1+e}} = 29.3\ \text{km/s}.$$

It is evident that velocity reaches its maximum at perihelion and minimum at aphelion. The

average velocity of the Earth's orbital motion equals half the sum of its maximum and minimum velocities:

$$v_{av} = \frac{2\pi a_{\oplus}}{T_{\oplus}\sqrt{1-e^2}} = 29.8 \text{ km/s}.$$

Now we shall consider the orbital motion of the Earth and planets differently. We shall not follow their travel round the orbits but focus upon the values which are conserved during motion. We know that the plane of revolution of a planet is constant and value A is conserved. These two statements can be integrated. Introduce vector **M** equal to the double product of A by the planet mass m and directed perpendicularly to the plane of revolution to the northern hemisphere of the celestial sphere. This vector is called the moment of momentum. The conservation of its value is equivalent to the second Kepler law, and its constant direction reflects the constancy of the plane of revolution of the planet.

The first two Kepler laws contain the principle of conservation of the moment of momentum. But these very laws manifest the conservation of one more vector. Note that the spatial orientation of the ellipse of planetary orbit and its eccentricity are constant. This means that another constant vector can be introduced, which is equal in value to the eccentricity and is directed at the perihelion. We shall designate it by **e**. It is clear that vectors **M** and **e** are mutually perpendicular.

The moment of momentum of the Earth by definition is perpendicular to the ecliptic and directed at point P, the center of the diagram of constellations in Fig. 2. The Earth passes its perihelion each year on January 3 and 4; in other words, vector **e** is directed at Gemini. The inaccurate dating does not mean a change in this direction, it is merely a specific feature of our calendar.

The laws of planetary motion were established for the Earth and five easily observed planets. Naturally, they proved to be true also for distant planets and minor bodies of the solar system. The first planet previously unknown to ancient astronomers was discovered in 1781 by the English astronomer W. Herschel. It was called Uranus (the astronomical symbol for it is ⛢ or ♅). In 1801 the Italian G. Piazzi discovered the first of the minor planets on the orbit between Mars and Jupiter. Since then orbits of more than 2300 minor planets, or asteroids*, have been determined. Many of them, in contrast to most planets, have a significant eccentricity and the angle of inclination to the ecliptic. In 1846 the French scientist U. Le Verrier and the English scientist J. Adams predicted the existence of one more planet by the disturbance of Uranus's orbit. That very year the German astronomer J. Galle discovered a planet just in the sector of the sky indicated by Le Verrier. That was Neptune designated by astronomical symbol ♆, the trident of Poseidon, the Greek sea-god identified

* Asteroid means 'star like'. This, not very appropriate, name was given to minor celestial bodies because their proper dimensions cannot be distinguished by telescopes; they look as point-like sources of light which is the reflected one in contrast to the light emitted by stars.

with Neptune by the Romans. The same thing happened to the ninth planet, Pluto. The American scientist P. Lowell had calculated the orbit of the yet undiscovered planet and C. W. Tombaugh discovered Pluto in 1930, 14 years after Lowell's death. Lowell's initials coinciding with the first letters of the name of the Greek god of the underworld were taken as Pluto's astronomical symbol: ♇ . The average distance from Pluto to the Sun is longer than that of Neptune. However, between 1980 and 1997 Pluto will be located nearer to the Sun than Neptune. This is due to the fact that Pluto will pass the perihelion of its orbit—the eccentricity of which is rather great—at the beginning of 1989. Although at first glance the orbits of Pluto and Neptune cross one another, no collision between them is possible since their motion is so arranged that the distance between them can never be less than 0.6 of the average distance from Neptune to the Sun.

Have a look now at the conserved characteristics of planets given in Table 2. The inclination of an orbit means here the angle of inclination of the orbit's plane to the ecliptic or, which is the same, the angle between the moment of momentum of the planet and the moment of momentum of the Earth.

Note that the angles of inclination of orbits are not great (Pluto has the maximum, then comes Mercury) and their eccentricities are also much less than 1 (the exclusion being again the planets which are the farthest and the nearest to the Sun). These peculiarities gave life to hypotheses about specific origins of Pluto and Mercury. The

small in mass Pluto could have previously been Neptune's satellite. Mercury is sometimes regarded as a former satellite of Venus. The real history of these planets is hard to be traced. In any case these events, the significant changes in their orbits, have occurred very long ago, at the beginning of the evolution of the solar system. Nowadays the planetary orbits are almost stable.

But what does the word "almost" mean? We have just stated that the laws of conservation of the moment **M** and eccentricity **e** are accurate. This is not absolutely so. These laws are true when only two bodies interact and the planet is acted upon only by the gravitation of the star. Yet there are many planets in our planetary system all of which gravitate not only to the Sun but also to one another.

We are fortunate that the mass of the Sun exceeds those of planets by so much that their mutual attraction results only in minor disturbances in orbits. Without this factor Kepler would have failed to discover his relatively simple laws and they would never incite the mind of Newton... But most likely neither would have happened because civilization can hardly exist in an unstable planetary system: climatic fluctuations on such planets are too great (we shall deal further with the reason therefore). The motion of even three bodies of comparable mass becomes extremely complicated and nonperiodical. The result may be that the lightest body would take energy from two others and leave the system for good. The solar system is stable because it has only one star, only one body the mass of which by far exceeds those of the others.

4*

Estimate the relation between the forces of attraction of the planets and the Sun. The relative action of Jupiter, the most massive of planets, on the Earth does not exceed the value

$$\frac{F_{♃}}{F_{\odot}} < \frac{m_{♃}}{m_{\odot}} \left(\frac{a_{\oplus}}{a_{♃}-a_{\oplus}}\right)^2 \sim 5 \times 10^{-5}.$$

The relative disturbance in the Earth's orbit produced by Venus is of the same order of magnitude (the effect of other planets is still lower):

$$\frac{F_{♀}}{F_{\odot}} < \frac{m_{♀}}{m_{\odot}} \left(\frac{a_{\oplus}}{a_{\oplus}-a_{♀}}\right)^2 \sim 3 \times 10^{-5}.$$

Mutual disturbances in the motion of planets are maximum at the moments of their conjunction. This is the term for the planets positioned on one line with the Sun and from one side of it. As this happens, the distances between the planets are minimum. The conjunctions of planets repeat periodically.

Let us find the period of conjunction of the planets with respective periods of revolution T_1 and T_2. The reciprocal of period is frequency. The revolution of all the planets takes place in one direction with the frequency which is the higher the nearer the planet is to the Sun. It is easy to see that the frequency of conjunctions equals the difference in frequencies of the revolutions of planets. Therefore the formula for the period of conjunctions T_{12} will be:

$$T_{12}^{-1} = |T_1^{-1} - T_2^{-1}|, \text{ or } T_{12} = \frac{T_1 T_2}{|T_2 - T_1|}.$$

Note this formula. We shall make use of it to calculate the periods of repetition of various

events on revolving planets and to understand the arrangement of our calendar.

Now we can estimate the specific times during which planets significantly change the orbits of one another, the times of violating the laws of conservation of moments and eccentricities of individual planets. After each conjunction the relative distortion of the orbit equals, by the order of magnitude, the ratio between the forces of attraction to the nearest planet and to the Sun. Consequently, the specific time during which the orbit of one planet would be significantly distorted by the gravitation of the other planet is

$$T_{12} F_{\odot}/F_{\mathrm{pl}} \sim T_{12} m_{\odot} (a_2 - a_1)^2 / m_2 a_1^2.$$

The time during which the Earth's orbit would be distorted under the influence of Jupiter, Venus, and, to a lesser extent, other planets is about twenty thousand years if calculated by this method. This is, however, only an estimate. Accurate calculations demonstrate that the eccentricity of our orbit changes nonperiodically but with the specific time equal to hundred thousand years (Fig. 11). The values of the Earth's eccentricity fluctuate about 0.028. This eccentricity is, presently, less than the average and continues to diminish. In 25 thousand years the Earth's orbit will become almost circular. The maximum eccentricity of the Earth's orbit reaches 0.0658 which is more than the present value by a factor of four. On the face of it, this value does not seem great since the orbit looks very much like a circular one with such an eccentricity. However, it turns out that the nonuniformity of

the Earth's orbital motion and the changes in the distance to the Sun influence significantly the thermal equilibrium and the climate of the Earth during a year and at great eccentricities. We shall deal with this fact in the closing section.

The moments of momentum of planets change under the action of disturbances with the same

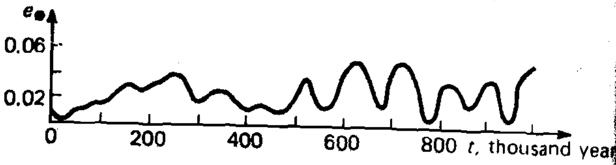


Fig. 11. Changes in the eccentricity of the Earth's orbit in the next million years.

specific times as the eccentricities. Yet they change not in the absolute value but only in direction. This is how it happens. We have already mentioned that the moments of momentum of all planets are almost parallel. The directions are contained in a narrow solid angle around the direction of the complete moment of momentum of the solar system. This moment is the vectorial sum of momenta of individual planets. It is a vector truly conserved in all planetary interactions. The value of the complete moment of momentum is 3.154×10^{43} kg·m²/s. This is a characteristic of the solar system exactly as fundamental as its mass equal to 1.993×10^{30} kg. The direction of the complete moment of momentum is also constant. This vector is directed at Draco. In the diagram of the celestial sphere given in Fig. 2 the point at which the vector is

directed is designated by letter E_0. The Earth's orbital moment perpendicular to the ecliptic is inclined to the moment of momentum of the solar system. This slope equals merely 1.7°. It is directed at Draco as well, at the center of Fig. 2 indicated by letter E. This point is called the pole of ecliptic.

The moments of momentum of all planets turn slowly around the complete moment of the solar system under the action of mutual disturbances. This phenomenon—the turn of the planes of planetary orbits—is called precession. The pole of ecliptic, which is the projection of the Earth's orbital moment onto the celestial sphere, would describe under the action of gravitation of other planets in a hundred thousand years a small circle with radius 1.7° around point E_0. Both this circle and the direction of precession are also indicated in Fig. 2. Frankly speaking, the trajectory of the pole of ecliptic is not strictly a circle since its radius also changes slowly.

The period of a hundred thousand years may seem extremely long, but nevertheless the period of precession is short compared to the time of existence of the solar system. This time, exceeding four billion years, can house at least forty thousand periods of precession. Certainly we cannot claim that the angles of precession and the angles of inclination of planetary orbits remain the same through such a great number of periods. They are most likely diminishing which is the reason for the solar system to have turned so flat.

If you have another look at Fig. 7, you would see that the very moment of momentum of the

solar system, which we have just considered as absolutely motionless, is in no way perpendicular to the galactic plane. This, in turn, means that it should also turn very slowly, i.e. precess. Yet, more or less accurate astronomical observations have been carried out for mere four hundred years and this time is insufficient for determining the actual mechanism of this turn. However, the characteristic time, the period of this precession, can be estimated. Since the Galaxy is significantly flattened and its mass is not concentrated in one nucleus but is distributed over the entire disk, the period of precession of the moment of the solar system should approach, by the order of magnitude, the Sun's period of revolution around the galactic nucleus, i.e. 2×10^8 years. Yet this period is notably less than the existence of the solar system: the Sun and the stars around it have accomplished between 10 and 30 turns around the nucleus of the Galaxy. But the moment of momentum of the solar system has not yet taken a stable position approaching a perpendicular to the galactic plane.

So this is the intricate route which the center of masses of the Earth and Moon takes in space: it moves around the Sun in an elliptic orbit with a period of one year; the alteration period of the ellipse eccentricity is about 10^5 years and the plane of the ellipse turns around a certain medium plane. The medium plane is perpendicular to the moment of momentum of the entire solar system and turns with specific time of 2×10^8 years. In addition, the Earth itself rotates. So let us take a closer look at the rotation of planets.

Chapter 2

Mobile Firmament of Heavenly Bodies

1. Kinematics of the Earth's Rotation

Imagine yourself to be an extraterrestrial approaching our planet from space. A sphere half illuminated by the Sun is facing you. Its surface is partially hidden by white clouds. Blue oceans and continents painted green by forests and fields, yellow by deserts, and white by snow look out here and there through the azure haze of the atmosphere and gaps between clouds.

An observer placed in space would soon notice that the motion of clouds is more or less random while the outlines of continents are stable. From the changing view of continents he would easily find out that the globe rotates as a solid body, i.e. as an integral whole and with one and the same angular velocity. It would not take the visitor from space a long time to determine the physical constants of our planet: its form is close to a sphere with the radius $R_\oplus = 6370$ km; the Earth's mass is $m_\oplus = 5.976 \times 10^{24}$ kg; the angular velocity of rotation is $\omega_\oplus = 7.292115 \times 10^{-5}$ s^{-1}. The outside observer would certainly measure the angular velocity of rotation in a star-oriented frame of reference. Note that the respective period $P_\oplus = 2\pi/\omega_\oplus = 86\,164.09$ s differs slightly from the period of 86 400 s which the terrestrials call the day. The cause of this difference will be treated below.

The axis of the Earth's rotation is directed at Ursa Minor (the corresponding point P is indicated near the North Star in the star diagram given in Fig. 2). The angle ε between the axis of the Earth's rotation and the direction to the pole of ecliptic E is 23°27′. This takes care of the extraterrestrial's mission. Let us consider now the human experience.

The frame of reference generally accepted on the Earth's surface is based on the axis of rotation. The well-known geographical coordinates are the latitude and the longitude. The axis of rotation crosses the Earth's surface at the points called poles; a circle of zero latitude, the equator, is positioned at equal distances from both poles.

As you know, the change of seasons is caused by the inclination of the Earth's axis. Consider Fig. 12. It is winter in the northern hemisphere, the month is December. The radius vector Sun-Earth is directed at Gemini. The Sun is seen from the Earth in the opposite Zodiac constellation, Sagittarius. A part of the Earth's surface receives no solar rays at all. The polar night sheathes the area within the polar circle while the Sun never sets near the South Pole.

The axis of rotation does not change its direction in the course of the planet's orbital motion. The areas of the polar night in the north and the polar day in the south gradually diminish. March 21 is the moment when the axis of rotation becomes perpendicular to the direction towards the Sun. This is the day of vernal equinox when day and night are of equal length on the entire planet. The Sun is in Pisces and the Earth is in Virgo if observed from the Sun. However, the

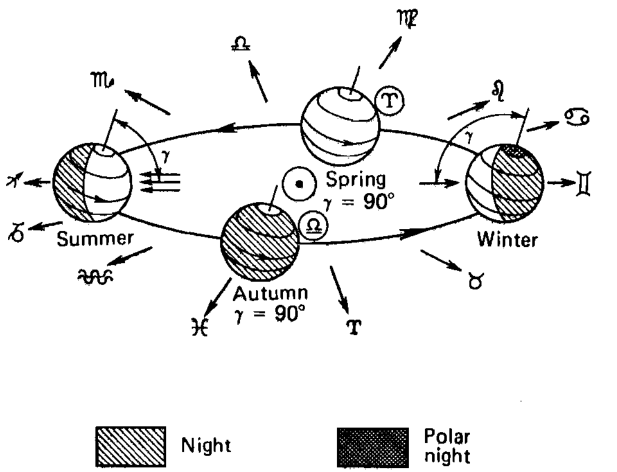


Fig. 12. Seasons of the year. Directions to Zodiac constellations are indicated by arrows.

point of vernal equinox is marked in Figs. 1 and 12 by sign ♈, the symbol of Aries. This is a tradition. More than two thousand years ago this point was actually located in Aries. We shall soon see why it has changed position.

Then summer comes to the northern hemisphere. June 22 is the day of summer solstice. The term "solstice" stems from the fact that the points of sunrise and sunset on the horizon and Sun's altitude at noon do not almost change for several days adjacent to the days of summer and winter solstices. The latitudes in which the Sun at noon stands during these days at zenith are called tropics. "Tropos" means a "turn" in Greek. The Sun seems to turn over the tropics during solstices.

The northern tropic is called the Tropic of Cancer. This is also a tradition since the Sun is in fact in Gemini. The southern tropic is called the Tropic of Capricorn although on December 22 the Sun is positioned in Sagittarius.

On September 22 or 23 the Earth passes the last of characteristic points of its orbit, the point of autumnal equinox. The axis of rotation is again perpendicular to sun rays, the length of the day is again equal to that of the night.

If you count the number of days between the vernal and autumnal equinoxes you may be surprised by the result of 186 which is more than half a year. In contrast, the period between the autumnal and vernal equinoxes is less than half a year. The obvious reason for that is the elliptic character of the Earth's orbit. The Earth passes its perihelion in winter therefore it has to go a longer way in summer while the velocity of its orbital motion is lower than in winter.

You know that the principle of the conservation of the moment of momentum results in the invariability of the orbital plane of the planet. One more effect of this principle is that the direction of the axis of the Earth's rotation is constant. To distinguish between the orbital moment of momentum and the moment of momentum of the Earth, associated with the planet's rotation around its axis, we shall further call the latter the moment of the Earth's rotation. Similarly to the axis of rotation, it is directed at the North Star. Its value is hard to be calculated since the Earth is a massive sphere and components of its mass are located at various distances from the axis of rotation. Moreover, the density of the Earth's interior increases towards the center of the planet.

The moment of momentum of a point mass m located at the distance r from the axis of rotation equals $m\omega r^2$. Thus, the moment of momentum of the Earth's rotation can be roughly estimated with the help of the expression

$$M_{\mathrm{rot}} \sim m_{\oplus}\omega_{\oplus}R_{\oplus}^2 \sim 10^{34}\ \mathrm{kg\cdot m^2/s}.$$

Compare this to the Earth's orbital moment of momentum (cf. Table 2)—it is less by many orders of magnitude. For this reason the rotation of planets can be ignored when the total moment of momentum of the solar system is calculated.

The Earth's moment of rotation does not change which is quite natural for a constant value (Fig. 13). But if its direction were measured with a higher accuracy and new results were compared to those previously obtained, it would come out that the moment of rotation does turn. Almost without changing the value it rotates slowly around the pole of ecliptic and around the orbital momentum of the Earth. The angle of its rotation in a year equals mere 20″, but in twenty six thousand years the moment would describe a cone around the pole of ecliptic and return almost to the initial position.

We have met this physical phenomena already—it is precession. We shall further see which forces cause the precession of the axis of rotation. For the time being we are interested in its kinematics in relation to the Earth's rotation. Thus, the period of precession is $T_{\Upsilon} = 2.578 \times 10^4$ years, the angular velocity being $\omega_{\Upsilon} = 2\pi/T_{\Upsilon} = 7.72 \times 10^{-12}\ \mathrm{s^{-1}}$. This is the velocity with

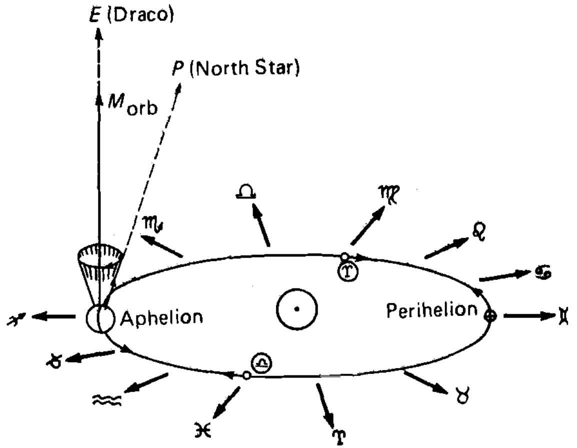


Fig. 13. Precession of the Earth's rotation. Diagram of the orbit with indicated points of vernal (♈) and autumnal (♎) equinoxes, perihelion ⊕, and directions of their motion.

which the point of vernal equinox ♈ travels round the ecliptic. Equinoxes occur somewhat earlier than the Earth makes a full turn in relation to stars. Therefore this phenomena is also called precession of equinoxes.

The precession of the axis of rotation was discovered by the great ancient Greek astronomer Hipparchus. In the 3rd century B.C. Greek astronomers made a list of stars having thereby compiled the first star catalog. 150 years later Hipparchus discovered that positions of all the stars had changed in relation to the vernal equinox. This means that the equinox comes somewhat earlier than a sidereal year passes.

Our calendar is so designed that the vernal equinox occurs annually on one and the same day: March 21. This fact permits us to calculate the time when the vernal equinox got its symbol of Aries and when the tropics of Cancer and Cap-

Table 3

Ancient and Modern Dates of the Sun's Passage Through Constellations

Constellation		Ancient dates	Modern dates
Aquarius	♒	January 20-February 18	February 21-March 18
Pisces	♓	February 19-March 20	March 19-April 21
Aries	♈	March 21-April 19	April 22-May 21
Taurus	♉	April 20-May 20	May 22-June 21
Gemini	♊	May 21-June 21	June 22-July 22
Cancer	♋	June 22-July 22	July 23-August 23
Leo	♌	July 23-August 22	August 24-September 23
Virgo	♍	August 23-September 22	September 24-October 24
Libra	♎	September 23-October 22	October 25-November 23
Scorpio	♏	October 23-November 21	November 24-December 23
Sagittarius	♐	November 22-December 21	December 24-January 22
Capricorn	♑	December 22-January 19	January 23-February 20

ricorn were named. The matter is that there still remains—beside these traditional terms—a quite commonly used correlation of Zodiac signs to months and dates. On these dates the Sun passed the boundaries of constellations in ancient times. Let us compare the dates of the Sun's passage through the constellations in ancient times and presently (cf. Table 3).

The table indicates that Cancer actually corresponded to the summer solstice as well as Capricorn did to the winter solstice. Therefore the names of tropics were fully justified. Nowadays this ancient correspondence between months and constellations has no real sense. Yet Zodiac signs are sometimes used to designate the moments of the vernal and autumnal equinoxes and the ancient dating is employed by astrologers, as also thousands years ago, to cast horoscopes.

The only thing the ancient correlation of dates and Zodiac constellations can be seriously used for is to determine the time of its introduction. We know the velocity with which the point of vernal equinox moves. In ancient times the Sun entered Aries on March 21 and in modern time it is located in Pisces on that day. The Sun covers the distance between the old and the new equinox points in 32 days. The period of precession equals approximately 26 000 years, during which time the point of equinox makes a full turn round the ecliptic and shifts by 365 days. Therefore a 32 days shift occurs in

$$26\,000\,\frac{32}{365}=2280 \text{ years.}$$

This means that the ancient correlation of dates to Zodiac constellations was accurate in about 300 B. C.

In 331 B.C. the army of Alexander the Great entered Babylon. In that raid Alexander was accompanied by many Greek scientists. Babylonian priests, the Chaldeans, were by far more advanced in astronomy than the Greeks. The Greeks borrowed from the Babylonians the system of units based on astronomical observations: the division of the day into 24 hours (two hours per each Zodiac constellation), the division of the circle into 360 degrees (the Sun passes about 1° of ecliptic each day), the division of the hour into 60 minutes and of the minute into 60 seconds, the division of the degree into angular minutes and seconds. It is thus not surprising that the Greeks designed the solar calendar in approximately 300 B.C. having taken the advantage of Chaldeans' astronomical knowledge. In that calendar months of the year strictly corresponded to Zodiac constellations and the Sun was really in Aries on the day of the vernal equinox.

It is easy to calculate the time when the Sun due to precession had displaced on the day of vernal equinox to Pisces, which made Pisces, in a certain sense, the major constellation of the Zodiac. However, it should be taken into consideration that in ancient times the boundaries between constellations were designated with lower accuracy than presently. Thus, we need only to estimate the moment of transition within a century: the result would be the 1st century B.C. This event, of course, was not a real astronomical phenomenon since the boundaries be-

tween the constellations are arbitrary. Nevertheless, it was not ignored by the ancients who attached so much importance to the positions of celestial bodies. There is a hypothesis that this very event had produced an increased activity in this sphere of mystical beliefs, the effect of which was the advent of Christianity. It is known, for example, that the religious symbol of early Christians was not the cross but a picture of a fish.

The next constellation after Pisces to house the Sun on the day of vernal equinox will be Aquarius. This transition will occur approximately in 2600.

2. Physical Background of the Calendar

The precession of the Earth's axis of rotation is a slow process quite negligible in a life-time of one generation. It is, however, the cause of astronomers' discomfort. They measure coordinates of stars from the point of vernal equinox, therefore they have to make a correction increasing with time in the longitude of stars, the correction first measured by Hipparchus. If a calendar is designed to serve more than one generation, the precession also should be taken into account.

As far as we, the residents of the Earth are concerned, the most important periodical events are the change of day and night and the change of seasons. Our everyday life is closely connected with the Sun. The change of seasons brings about weather changes on the planet and the agricultural production also depends upon weather. Thus

the calendar based on the periodic character of these phenomena is most convenient for us but to have matched it was not an easy task.

Let us again write down the adequately accurate physical data required to construct a solar calendar:

sidereal year	$T_{\oplus} = 3.155815014 \times 10^7$ s
sidereal day	$P_{\oplus} = 8.616409 \times 10^4$ s
period of precession	$T_{\Upsilon} = 25780 \times 3.156 \times 10^7 \text{ s} = 8.135 \times 10^{11}$ s.

Determine the period of time between two successive vernal equinoxes. This period is called a solar year (or a tropical year). It is this year on which a calendar should be based to make equinoxes and seasonal changes fall on the same dates.

The point of equinox on the celestial sphere is determined by the moment at which the direction from the Earth to the Sun is perpendicular to the axis of rotation. This point transits along the ecliptic in the direction reverse to the Earth's motion and with the period of precession T_{Υ}. Calculating the periods of planets' conjunctions, we have seen that the frequency of conjunctions is equal to the difference in the frequencies of rotations. A similar procedure can be employed to calculate the period between equinoxes. However, planets rotate in one direction and the point of the equinox transits opposite to the Earth's motion. For this reason, frequencies, the reverse periods of these motions, should be added to determine the reverse solar year:

$$T_{\text{sol}}^{-1} = T_{\oplus}^{-1} + T_{\Upsilon}^{-1}; \quad T_{\text{sol}} = \frac{T_{\oplus} T_{\Upsilon}}{T_{\oplus} + T_{\Upsilon}} = 31{,}556{,}926 \text{ s}.$$

The solar year is by 1224 seconds, or by twenty and a half minutes, shorter than the sidereal year, the true period of the Earth's revolution.

The above formula gives, in fact, the value of the circular motion period in a frame of reference rotating with another known period. It is also valid for calculating solar days. Since the Earth's orbit has a form of an ellipse, the Sun transits along the ecliptic nonuniformly over a year. Therefore the angular velocity of Sun's motion over the celestial sphere slightly changes. To construct a calendar, a mean solar day of the year, P_0, is to be found. In other words, we have to determine the period of the Earth's rotation in the frame of reference where the Sun-to-Earth vector does not rotate, on the average, during a year. Since the Earth's rotation and revolution occur in one direction, the mean frequency of the Sun's appearance P_0^{-1} is the difference of frequencies: the frequency of the Earth's rotation $P_{\oplus}^{-1}$ and the frequency of the occurrence of equinoxes T_{sol}^{-1}. The result is the following:

$$P_0^{-1} = P_{\oplus}^{-1} - T_{sol}^{-1}; \quad P_0 = \frac{P_{\oplus} T_{sol}}{T_{sol} - P_{\oplus}} = 86\,400.000 \text{ s.}$$

Here comes the familiar to us number of seconds in a day: $86\,400 = 24 \times 60 \times 60$. Hence the second was previously determined as 1/86 400th part of the mean solar day. However, in the 60s the atomic clock was invented whose accuracy is higher than the stability of the Earth's rotation. The standard of time was changed then and the second is presently determined by the frequency of oscillations of a cesium atom. This has also provided an opportunity to measure minor disturbances in the angular velocity of the Earth's rotation and extremely low deceleration of its motion.

Now let us calculate the number of solar days in a solar year:

$$N = \frac{T_{sol}}{P_0} = 365.2422 \ldots$$

The first three figures of this number are well familiar to everybody: this is the number of days in a year. If the fractional part is rounded off to $0.25 = 1/4$, the origin of leap-years becomes quite clear: one more day (February 29) should be added every four years to keep the date of equinox down to March 21.

However, the divergence of N and its approximation to 365.25 is still significant. A one-day shift of the equinox will occur in such a calendar in $(N - 365.25)^{-1} = 128$ years. For this reason the presently accepted calendar has not 100 but 97 leap-years per each 400 years. For example, 1600 was a leap-year, 1700, 1800, and 1900 were not, and 2000 will be again a leap-year, and so forth.

The approximation $365 + \frac{1}{4} - \frac{3}{400} = 365.2425$ also does not absolutely coincide with N. However, it is not so easy to calculate the moment when the date of vernal equinox would shift by one day. The matter is that number N itself is not strictly constant. It diminishes very slowly each year by mere 6×10^{-8} due to deceleration of the Earth's rotation. In any case, the modern calendar can be safely operative for the nearest thousand years without any corrections in dates. The vernal equinox will stick to March 21.

The humanity has adopted the decimal system for nearly all units of measurement with the exception of time and angles. The long life of complicated Babylonian units, connected by relationships, which may seem strange at first sight (360, 60, and 24), has its own reasons. First, any radical change in the calendar or units of time is very toilsome since all historical dates have to be recalculated. Second, the basic natural units of time, the year and the day, are incompatible with any units of time system. The adopted system of time and angle measurement has one advantage. The Sun transits round the ecliptic (Fig. 1) approximately by 1° per day. It is also useful to know that the Sun and the Moon transit over the celestial sphere by one half of their respective disks per minute.

3. Rotation of the Moon and Planets

All planets and their satellites rotate. The rotation of planets conforms to a certain regularity: the more the mass of the planet, the faster it rotates. It is not a strict and proved law, there are exclusions from this rule, but on the whole (cf. Fig. 14) this regularity looks to be true.

The answer to it may lie in the history of formation of the solar system. According to modern concepts the Sun and planets had formed out of a rotating nebula composed of gas and solid dust particles. The particles collided and joined into larger bodies thereby forming embryos of the Sun and planets. The largest number of collisions occurred at the center of mass of the system where almost all the gas of the nebula had been

attracted. So the Sun was formed. But almost all the initial moment of momentum of the nebula appeared to be concentrated not in the Sun but in planets. After the Sun was ignited (the mechanism of its burning is treated in Chapter III),

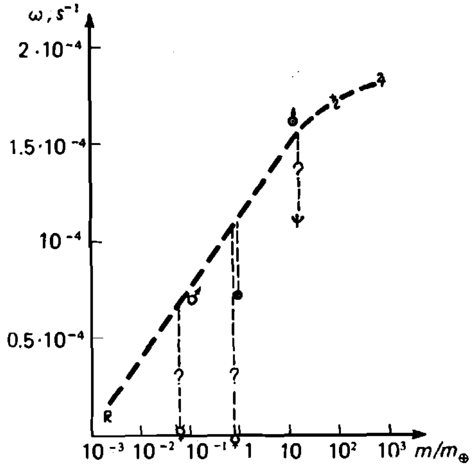


Fig. 14. Dependence of the angular velocity of rotation on the masses of planets. Corresponding points are indicated by symbols of planets, straight lines with question marks indicate a possible deceleration of rotation.

its radiation dispersed the light gases from the immediate surroundings to peripheral areas to form there giant planets. Planets of the terrestrial group turned out to be composed of solid particles.

The matter of the rotating nebula being compressed into dense spheres, the velocity of rotation increases due to the principle of the conser-

vation of the moment of momentum. Thus it is not surprising that the giant planets have a higher velocity of rotation than that of smaller planets.

It is interesting to have a look at deviations from this regularity represented in Fig. 14 by a line connecting the majority of planets. Let us assume that a similar regularity was operative with all planets in the course of the formation of the solar system but some planets decelerated their rotation later for different reasons. This probable evolution of rotation is indicated in Fig. 14 by dashed lines with question-marks.

Neptune is very far from us and our knowledge about it does not yet suffice to discuss seriously the "correctness" of its rotation. Recall however the assumption that Pluto had previously been its satellite. Maybe Neptune has partially lost the moment of rotation after Pluto had broken away?

We also know the cause of the deceleration of the Earth's rotation and the present value of this deceleration. The Earth is braked by its satellite, the Moon. The actual mechanism of this process will be treated below. When the evolution of the solar system had just started, *the deceleration was apparently reciprocal:* not only did the Moon brake the Earth's rotation but the Earth also impeded the fast rotation of its satellite. As the result of this the Moon is facing the Earth always by one side.

This does not imply that the Moon does not rotate in relation to stars, but its sidereal period of rotation equals exactly the Moon's period of

revolution around the Earth: $P_{☾} = T_{☾} =$ 27.322 days. The kinematics of Moon's rotation is clear from Fig. 15.

If the ratios of the frequencies of oscillations or the frequencies of rotations are multiple of the ratio of integers, the frequencies are said to be in

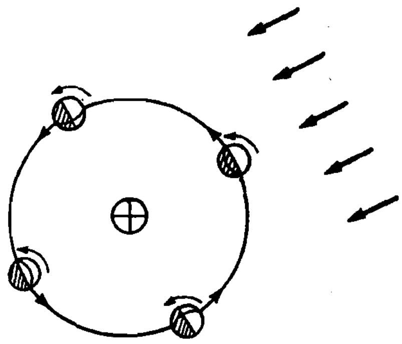


Fig. 15. Rotation of the Moon and succession of its phases. The crescent indicates Moon's side facing the Earth, the cross-hatching denotes the shady side of the Moon.

resonance. The term stems from the Latin word *resonare* meaning "to repeat sound". The importance of this physical phenomenon consists in that the resonance is indispensable for the interaction of bodies which supports the multiplicity of frequencies and makes the resonance stable. The Moon's rotation is resonant to its revolution, but the Moon is not the only body of the solar system to rotate resonantly.

The cause of the anomalously slow rotation of Mercury and Venus is not clear. The explana-

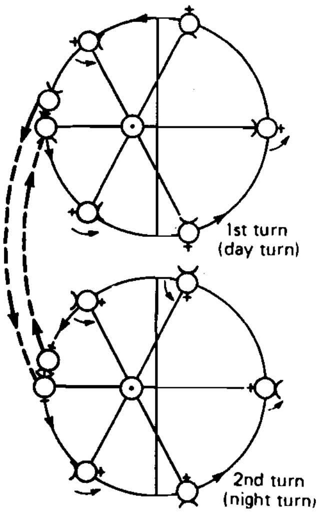


Fig. 16. Rotation and revolution of Mercury; the diagram shows two successive turns of the planet around the Sun: the "day turn" and the "night turn" in relation to Mercury's side marked by "horns" of its astronomical symbol.

tion is probably to be looked for in the hypothesis which reads that in the past they revolved around the Sun together in one orbit. It is interesting, however, that presently both planets take part in resonant rotation but their resonance is not between themselves.

Mercury's period of rotation was measured not very long ago with the help of radio astronomy. It turned out to be exactly equal to two thirds of the planet's period of revolution. The resonance results in a very special kinematics of Mercury's motion. Consider Fig. 16. The ellipse of Mercury's orbit has a significant eccentricity, therefore, the Sun positioned in the focus of the ellipse is notably displaced from the center. The velocity

of Mercury's transit round the orbit is rather nonuniform: it is by 1.52 times higher at perihelion than at aphelion.

Mercury's period of revolution $T_{\mercury}$ is 88 terrestrial days and its period of rotation is $P_{\mercury} = 2T_{\mercury}/3 = 58.7$ days. Let us employ the formula of the solar day to find $P_{0\mercury}$, the period between two sunrises on Mercury:

$$P_{0\mercury}^{-1} = P_{\mercury}^{-1} - T_{\mercury}^{-1} = \frac{1}{2} T_{\mercury}^{-1};$$

$$P_{0\mercury} = 2T_{\mercury} = 176 \text{ days.}$$

Thus Mercury's solar day is three times longer than its sidereal day and twice longer than its period of revolution. For this reason Mercury's orbit had to be sketched twice in Fig. 16: once for a day turn and once for a night turn. Hornlets and cross-tail of the astronomical symbol serve to facilitate the description of its rotation. Note that in perihelion Mercury is turned to the Sun either by one or other side while in aphelion it faces the Sun either by horns or the tail.

The rotation of Venus is still slower but it occurs in reverse direction! All the planets move in one direction around the Sun, the majority of them and their satellites rotating in the same direction. This means that their moment of rotation approaches the direction of the orbital moment of momentum, the exclusion being Venus and Uranus. The axis of Uranus's rotation is almost perpendicular to its orbital moment of momentum and is almost in the plane of ecliptic.

The axis of Venus's rotation is almost perpendicular to the plane of its orbit but the direction of its rotation is reverse. Therefore the value of its period of rotation should be written with a minus: $P_{♀} = -243.16$ days.

The resonance of Venus's rotation turned out to be related to the Earth's orbital motion. The period of Venus's rotation $P_{♀}$, its period of revolution $T_{♀}$, and the Earth's period of revolution make up the following accurate equality:

$$P_{♀}^{-1} = -4T_{♀}^{-1} + 5T_{\oplus}^{-1}.$$

To understand the kinematics of Venus's motion, the period of the Earth and Venus conjunctions is to be calculated:

$$T_c^{-1} = T_{♀}^{-1} - T_{\oplus}^{-1}; \quad T_c = \frac{T_{♀}T_{\oplus}}{T_{\oplus} - T_{♀}} = 583.92 \text{ days}.$$

With such a period these planets approach one another at a minimum distance. Consider now a rotating frame of reference in which the Earth is motionless. In that case T_c would be a time during which Venus returns to the Earth after having made one more turn around the Sun. It is easy to check that the period of conjunction, i.e. 583.92 terrestrial days, covers exactly five solar days on Venus which means that the Sun would rise five times above Venus's horizon during that time. An observer placed on Venus would see four "earthrises" during the same period if the planet's atmosphere were transparent:

$$\frac{T_c}{5} = (T_{♀}^{-1} - P_{♀}^{-1})^{-1}$$

$= 116.8$ days (a solar day of Venus);

$$\frac{T_c}{4} = (T_{\oplus}^{-1} - P_{♀}^{-1})^{-1}$$

$= 146.0$ days (a "terrestrial" day of Venus).

At the moments of conjunction, when Venus is nearest to the Earth, it always faces us with one

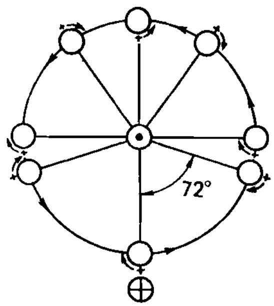


Fig. 17. Rotation of Venus and its revolution around the Sun in a system of coordinates of a motionless Earth. At the moments of conjunction Venus always faces the Earth by the side marked by the cross of its astronomical symbol.

and the same section of its surface. This place is marked in Fig. 17 by the cross of Venus's astronomical symbol.

Thus, planets that do not conform to the regularity indicated in Fig. 14 have their special characteristics of rotation. This makes us believe

that the rule "a faster rotation to a larger planet" is not accidental.

Note one more significant difference between the rotation of terrestrial-type planets and giant planets. Mercury, Venus, the Earth, and Mars rotate as a whole, as a solid body. The velocity of rotation of giant planets is not homogeneous. It is notably dependent on latitude, and, apparently, on depth. (The angular velocities given in Fig. 14 are calculated for the periods of rotation of equatorial areas.) This indicates that giant planets are not solid bodies.

We have so far discussed only the kinematics of the planetary rotation while internal forces of planets induced by rotation and gravitation have been ignored. Further we shall treat the internal dynamics of planets.

4. Forms of Celestial Bodies

The great Russian writer Leo Tolstoy said that all happy families are alike while each unhappy family is miserable in its own way. The same can be said about planets: all planets and their large satellites are alike while each asteroid and a small satellite are shapeless in their own way. The form of the Sun and planets is a sphere slightly flattened by rotation. This is so because the sphere is a stable form of a heavy body. But where is the boundary between minor and major celestial bodies? What is the critical mass and size to gain a spheric form?

Celestial bodies become spherical under the action of their own gravity. A small asteroid may retain an arbitrary form, but there is a certain

limiting mass at which the acceleration of free fall becomes so great on the body's surface that the rock cannot sustain its own weight: projecting parts of the asteroid will break away to yield a spherical body. Let us try to estimate the mass and size of a celestial body still able of taking up an arbitrary form.

Let the mass of a celestial body be m and its average density ρ. In this case its volume is m/ρ and the size of the body approaches $R \sim (m/\rho)^{1/3}$. Our objective is not an accurate result but an estimate. An estimate may ignore numerical coefficients not very much different from unity. So let us consider R as the specific size of a body, regardless of its shape for the time being.

The free fall acceleration on the surface of a celestial body can be estimated as $g \sim Gm/R^2 \sim Gm^{1/3}\rho^{2/3}$. It is directed approximately to the body's center of mass. Estimate now the pressure produced in its central part under the action of its own weight. As you know, the pressure produced by a homogeneous column with height R in the field of gravity g is $p = \rho g R$. In our case the acceleration changes along the radius and this formula becomes inaccurate, but quite applicable for an estimate. Thus, pressure inside a celestial body equals, by the order of magnitude,

$$p \sim \rho g R \sim G m^{2/3} \rho^{4/3}.$$

Pressure compressing a material from all sides does not destroy it. However, if the body's shape is irregular, i.e. nonspherical, shear stresses are induced in it by gravity. Stress is a physical term describing the distribution of internal forces

in a material. Pressure is also a stress but such that the forces applied to each point of a body act on the given point uniformly from all sides, i.e. isotropically. In addition to the pressure, the tension of the uniform compression (Fig. 18*a*),

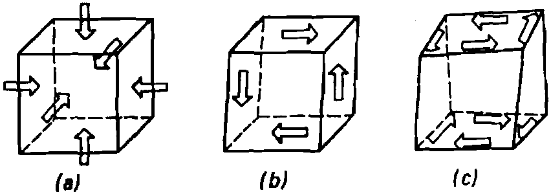


Fig. 18. Stresses in a solid body: (*a*) pressure or compression from all sides; (*b*) shear and pressure; (*c*) simple shear or torsion.

a shear stress is also possible. The latter occurs in a situation when the forces are exerted as indicated in Fig. 18*b*, *c*.

In liquids or gases shear stresses produce currents. Therefore the familiar to you Pascal's principle holds true for liquids and gases at rest since tensions are isotropic in such a medium. Solid bodies elastically resist shear stress. This is why the existence of arbitrary shaped solid bodies is possible, which concerns not only celestial bodies since there are shear stresses almost in any solid body. If the form of a celestial body is other than the sphere, shear stresses in it are of the same order of magnitude as pressure: $\sigma \sim p \sim Gm^{2/3}\rho^{4/3}$.

Solid bodies resist shear stress elastically to some limit, which depends on the material. This limit is either the shear strength for brittle bodies

which crack or crumble, or the yield strength for plastic materials. In either case, there is an ultimate strength for all solids. Let us designate it by σ_s. When the stress exceeds σ_s the body changes its form irreversibly.

Therefore the condition $\sigma \sim \sigma_s$ provides an estimate for the critical mass m_{cr} and critical size R_{cr} above which celestial bodies may only have a shape approaching the sphere. Here are the formulas for the critical mass and size:

$$m_{cr} \sim \frac{1}{\rho^2}\left(\frac{\sigma_s}{G}\right)^{3/2}; \qquad R_{cr} \sim \frac{1}{\rho}\left(\frac{\sigma_s}{G}\right)^{1/2}.$$

Estimates for some cosmic and terrestrial materials are given in Table 4.

Table 4

Maximum Masses and Dimensions of Shapeless Solid Celestial Bodies

	Ice	Lunar rock	Granite	Iron
Density (10^3 kg/m^3)	1.0	2.5	2.7	7.8
Ultimate strength (N/m^2)	3×10^6	3×10^7	10^8	10^9
Mass m_{cr} (kg)	10^{19}	3×10^{19}	3×10^{20}	10^{21}
Radius R_{cr} (km)	200	300	500	500

The dimensions of all the planets of the solar system are by far larger than critical and they all are, in fact, spheres. Yet Phobos, the satellite of Mars, looks like a potato sized 14 × 11.5 × 10 km which is less than critical. Amalthea,

a moon of Jupiter, is 265 km long and mere 150 kn wide. One more example of a body of near-critica size is given by Mimas, a moon of Saturn. Thi is a sphere 390 km in diameter with mass 3 × 10^{19} kg and mean density approaching tha of water ice. Yet its spherical surface is distorte by a deep crater 130 km in diameter which i a probable result of some other body's fall.

It is clear that on celestial bodies with mas approaching the critical, height of mountain should be comparable to the radius of the body And what is the ultimate height of mountain on large solid planets?

A steep mountain is also subject to shear stresses Let a mountain with density ρ be a cone o height h. In this case the mean pressure on it base equals $p = \rho gh/3$. If steepness exceeds 45 shear stress in some areas inside the mountai reaches pressure p. If the latter is related to th ultimate strength σ_s, we get a formula to esti mate the maximum height of mountains o planets:

$$h_{max} \sim \frac{3\sigma_s}{\rho g}.$$

The ultimate height of the Earth's granit mountains assessed by this formula is 11 km Assuming that the rocks on other planets hav the same characteristics, we get that Venu cannot accommodate a mountain higher tha 13 km, Mars and Mercury dictate a 30 km limit

Compare the above estimate with the actua data about the highest points of planets: th Earth's Everest is 9 km high, Maxwell volcan on Venus is 12 km high; Arsia Mons and Olympu

Mons volcanoes are the highest mountains on Mars with the height respectively 27 and 24 km.

You see that surface deflections from the average depend essentially on the free fall acceleration on planets, the height of mountains being insignificant compared to their radii. Therefore, the

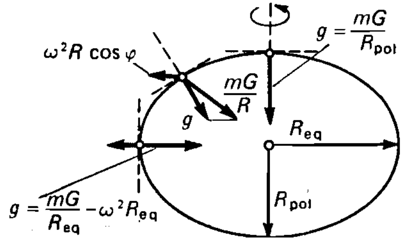


Fig. 19. Equilibrium form of a rotating planet. Normal to the surface is directed along the vector sum of gravitational and centrifugal acceleration.

shapes of giant planets as well as of solid terrestrial-type bodies approach their stable, equilibrium state. Consider now the effects of rotation on the shapes of planets.

If a planet of size R, which is larger than critical, rotates, its equilibrium form is not an ideal sphere but a sphere flattened at the poles. The radius of equator R_{eq} of the rotating planet is greater than the distance from the center to the pole R_{pol}. Figure 19 indicates that a stable shape of a planet is characterized by the perpendicularity of its surface to the vector sum of gravitational and centrifugal accelerations at any point. The ratio of these values $\omega^2 R/g = \omega^2 R^3/mG$ is a major factor determining the

shape of a rotating planet. It turns out that a planet's oblateness, i.e. a relative deviation of its surface from the spheric form, approaches the ratio of accelerations

$$\frac{R_{eq}-R_{pol}}{R_{eq}} \approx \frac{\omega^2 R^3}{mG}.$$

Check this statement against Table 5. It is clear that this primitive formula provides an acceptable agreement between estimates and actual flattening of planets.

Table 5

Periods of Rotation and Oblateness of Planets

	R (m)	P (s)	$\frac{\omega^2 R^3}{mG}$	$\frac{R_{eq}-R_{pol}}{R_{eq}}$
Earth	6.37×10^6	86 164	1/289	1/298
Mars	3.39×10^6	88 643	1/220	1/190
Jupiter	7.14×10^7	35 430	1/11	1/15
Saturn	6.03×10^7	36 840	1/6	1/9.5

According to the accurate theory of liquid sphere equilibrium the Earth's compression is 1/299.67 while its actual oblateness equals 1/298.26. This exceeds the calculated value, therefore it corresponds to a higher velocity of the Earth's rotation: 1.002 $\omega_\oplus$. This was the velocity of the Earth's rotation 10 million years ago. In such a long time the internal terrestrial rocks "flow over" to a new equilibrium state. This is the time it takes the Earth to tune its shape to rotation.

Thus the Earth approaches the equilibrium shape. Amazingly, the Earth's surface is divided

into continents and oceans with distinct borders between them. The continents rise above the ocean bottom by 4 to 5 kilometers. An individual mountain that high would not surprise anybody but the area of continents is enormous: 30% of the Earth's entire surface. Though oceanic depressions are filled with water but water density is twice or thrice lower than that of rocks, therefore water cannot compensate the pressure they produce by all these five kilometers of solidity. Why do the continents fail to submerse then? Terrestrial rocks are sufficiently pliable. Assumingly, the Earth's surface should sooner or later become smoother but for some reason this does not happen. Is there any explanation for that?

The matter is that continents actually float. The Earth's crust is composed of two major types of rock: basalts and granites. These are covered by a layer of sedimentary rocks with a lower density. The density of basalts is higher than that of granites at similar temperature and pressure. Five-kilometer thick continental mass floats on granite "cushions" the thickness of which varies between fifteen and twenty kilometers. The granite and basalt are underlied by the Earth's mantle composed of even denser rocks. The actual composition of the mantle is so far unknown since we are still failing to drill the Earth's crust that deep.

The continents float but only from the hydrostatic point of view. They also transit in relation to one another together with adjoining areas of oceanic crust. Continents move by several centimeters per year. The striking similarity of African and South-American coastlines is not acci-

dental: about 200 million years ago these continents were a unified whole. The opposite coasts of Atlantic Ocean, which is in fact a widening gap between these continents, also have similar geological structures.

About 100 million years ago Australia was integrated with Antarctica. The center of that parent continent was positioned, however, not at the South Pole but approximately in latitude 60° South. Bones of the ancestors of marsupials, which now live only in Australia, were found in Antarctica.

The hypothesis of continental drift was originated by a German geophysicist A. Wegener as early as 1912. For many years the hypothesis was considered groundless. It revived only in the 60s but as a theory supported by various evidence about the structure of the ocean floor. This theory is called global tectonics, or tectonics of plates. The basic principle of this theory reads that it is not the continents that move but plates, large areas of the Earth's crust including both continents and adjoining sections of the ocean floor. There are six major plates: Euroasiatic, African, Antarctic, Indo-Australian, American, and Pacific. Several minor plates are positioned between them and move to some extent independently.

At some boundaries of plates a new crust is being formed. All these boundaries are situated in oceans. For example, the mid-Atlantic ridge runs the length of the Atlantic Ocean. The Earth's crust builds up towards American plate on the one side and towards African and Euroasiatic plates on the other side. There are boundaries where plates collide. In this case one of them submerges under the other, which is happening in Far East where the Pacific plate submerges under the Euroasiatic plate.

The new crust's birth being a relatively quiet event, the old crust's burial is attended by heavy earthquakes and volcanic eruptions. Earthquakes are caused by friction on the boundaries of the plates moving together. Accumulated shear stresses exceed periodically the ultimate strength of rocks and this results in powerful tremors of the Earth's crust. Volcanic eruptions result from the heating of the sedimentary layer of the submerging plate. The heating originates chemical reactions and gaseous products, mainly water vapor and carbon dioxide, rise to the surface. The outlet of gases forms a volcanic chain along the boundary of colliding plates.

Unfortunately, this book does not have all that space required for a detailed discussion of these, most interesting, natural phenomena. Thus, we shall limit ourselves only to the source of energy of tectonic phenomena.

Geological processes are accompanied by differentiation, a separation of matter in the gravitational field: rocks with higher density mainly descend while the lighter move upwards. Thus, forces applied to blocks of individual rocks are identical to Archimedean forces in liquids, while the energy, which is eventually responsible for the mobility of the Earth's material, is a potential energy of masses with different chemical composition in the field of gravity. This energy is released in the form of heat in the course of tectonic phenomena.

The actual causes of the continental drift, i.e. the motion of plates, are not sufficiently clear yet. It is however interesting that tectonic activity—the volcanic eruptions—has also been discovered on some other planets and satellites. This proves that similar processes take place in the interior of those bodies.

Really, the celestial bodies resemble living creatures to some extent but on the cosmic scale of time and space. Hundreds million years pass and the surface of planets and satellites changes reflecting thereby conversion of the chemical and physical structure of their interior. Minor celestial bodies seem to die after having exhausted their store of energy and cooled down, the life of larger bodies is extremely long. A special complexity of the geological life of the Earth, duration of its evolution, is related to its massive and relatively close satellite, the Moon.

5. The Earth and the Moon

This planet has only one natural satellite. The Moon, however, is a unique phenomenon for the solar system. There are no other examples of so great an effect of a satellite upon a planet as that of the Moon upon the Earth. In this section we deal with that very influence, how it manifests itself and the mechanism of interaction.

The ratio of the mass of the Moon to that of the Earth is $m_☾/m_\oplus = 1/81.3$. Only Pluto has a moon called Charon with a probably higher ratio of respective masses but, apparently, the evolution of this planetary system has already ceased since both bodies face one another always by one side. Other planets have moons the masses of which do not exceed a thousandth fraction of the mass of the mother-planet.

The average distance to the Moon, $a_☾ = 3.84 \times 10^8$ m, is by 60 times longer than the Earth's radius. There are satellites in the solar system which revolve relatively closer to their planets, but nowhere except the Earth such beautiful solar eclipses are observed. You certainly know the cause of eclipses. A specific feature of eclipses observed on the Earth is that the angular diameters of the Sun and the Moon are approximately equal. These diameters are not constant due to the ellipticity of the Earth's orbit around the Sun and that of the Moon around the Earth ($e_☾ = 0.055$). The Sun's diameter varies between 31.5′ and 32.5′ while that of the Moon is between 29.4′ and 33.5′. This "accidental" coincidence is realized exclusively on the Earth due to which fact we can observe both total solar eclipses, when the Moon's disk covers the Sun, and annular eclipses, when the lunar subtense is less than the solar subtense and a ring of light from the Sun is visible. A solar eclipse is a rare event for each given area but on the whole 43 solar eclipses are observed on Earth during 18 years.

An eclipse is a physically simple phenomenon. So let us wave good-bye to its elegance and proceed to the Moon's dynamic effects which are of much greater interest to us.

The force of attraction between the Earth and its satellite is $Gm_\oplus m_☾/a_☾^2$. It causes the Moon's orbital motion with the period $T_☾ = 27.3$ days. A force of equal magnitude acts upon the Earth. The Earth revolves with the same period around

the center of mass of the Earth-Moon system. When we discussed the Earth's orbit, it was the orbit of the center of mass, or the Earth's orbit averaged over monthly variations, which was actually meant. It is easy to find out that the amplitude of the Earth's deviations from this average orbit caused by the lunar attraction equals $a_☾ m_☾/m_⊕ = 4700$ km. The trajectory of this planet around the Sun is slightly corrugated. Each full moon, when the Sun and the Moon are positioned at the opposite sides of the Earth, we are by 1.5 terrestrial radii closer to the Sun than at the nearest new moon. These monthly variations are however much less than the yearly variations in the distance to the Sun caused by the eccentricity of the Earth's orbit. They are important for the accurate astronomical observations of the solar system bodies. Other significant physical effects of the Earth's corrugated motion are unknown. Yet the interaction between the Moon's motion and the Earth's rotation yields several effects significant both for their daily manifestation and for the history of our planet.

Let us start with a qualitative explanation of the precession of the axis of rotation. Already Newton knew that the Moon was the major cause of precession. The precession results from the nonuniform gravitation of the equatorial broadening of the Earth's shape to the Moon.

The lunar orbit can be considered elliptic only in a rough approximation. The plane of the lunar orbit and the direction to perihelion change rather fast with periods of several decades due to the action of the Sun and planets. As you remember, the period of precession is by far longer. Thus, to

understand the cause of precession and the actual position of the Moon at any given moment is insignificant. The important thing is how often in these twenty six thousand years does the Moon appear at points close to the modern orbit, what does the "doughnut", i.e. the torus, formed by

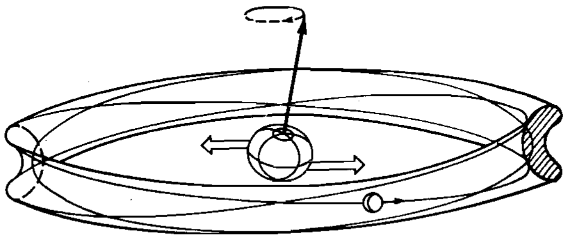


Fig. 20. Torus of possible lunar orbits and forces causing the Earth's precession.

a multitude of lunar orbits during that period, look like (Fig. 20). The external radius of this torus approaches the distance to the apogee $a_☾(1 + e_☾)$ while the internal radius approximates the distance to the perigee $a_☾(1 - e_☾)$. The thickness of this doughnut is approximately equal to $2a_☾ \tan i$, where $i = 5°9'$ which is the mean inclination of the lunar orbit to the plane of ecliptic. The doughnut of the aggregate lunar orbits is positioned parallel to the plane of ecliptic, since orbital alterations are caused by the Sun and planets.

If we assume for a moment that the mass of the Moon is distributed all over the doughnut's volume, its effect on the Earth would equal that of the Moon averaged through a long time. It is

then easier to see the direction of this actio
The Earth is just slightly flattened, by mere o
three-hundredth, but it is sufficient for the attra
tion by the doughnut of the equatorial broadenin
of the Earth's shape to produce a pair of force
i.e. the moment of forces. It tries to turn th
Earth so that the equator would coincide wit
the plane of ecliptic and the axis of rotatio
would coincide with that of the doughnut of luna
orbits.

The laws of mechanics of a rotating body a
rather complicated. You know that a whirle
top behaves quite differently from a nonrotatin
top: instead of falling aside it persists in rota
tion while the top's axis of rotation precesse
i.e. describes a cone. Just in the same way th
Earth's axis does not *immediately* take a stab
position. It precesses around the perpendicula
to the ecliptic. Those who are familiar with th
dynamics of a rotating body can easily translat
all these speculations into the language of nu
merical estimations and work out an expressio
of the Earth's period of precession:

$$T_{\Upsilon} \sim P_{\oplus}\frac{m_{\oplus}}{m_{\leftmoon}}\left(\frac{a_{\leftmoon}}{R_{\oplus}}\right)^3 \sim 5\times 10^4 \text{ yeart.}$$

The assessment agrees rather well with the tru
value.

If there were no Moon, the axis of rotation woul
precess just the same under the similar action o
the Sun and planets but the period of precessio
in such a case would be about a hundred thou
sand years. The total action of the solar syste
causes not only a precession with a constant angl
of inclination but also a slow change in th

very angle of inclination ε of the axis of rotation to the ecliptic. The results of accurate calculations of the future alterations of value ε are given in Fig. 21. The character of past alterations of the angle of inclination ε was the same: the major period of its variations is 41 thousand years, while that of the envelope of amplitudes is two hundred thousand years. These minor changes in

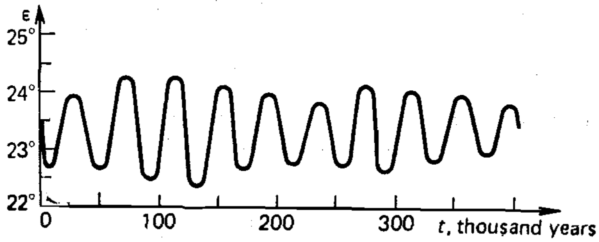


Fig. 21. Variations of the slope of the axis of the Earth in the future 400 thousand years.

the inclination of the axis of rotation to the plane of ecliptic are responsible for climatic alterations on the Earth which we shall discuss in the final section of this book.

The most apparent consequence of the close vicinity of a large satellite is certainly not precession but tides and ebbs. Twice a day, actually every 12 hours 25 minutes, the level of water in the seas rises by approximately one meter, and a quarter of a day later the sea retires. Tides and ebbs are an obvious demonstration of minor distortions of the Earth's shape under the influence of the lunar attraction. Why then is the major period of tides equal to approximately

half a day? Each point of the Earth in the cours of rotation approaches the Moon with a perio

$$(P_{\oplus}^{-1}-T_{☾}^{-1})^{-1}=24 \text{ hours } 50 \text{ minutes}$$

which is twice longer than that of tides. So wh is the cause of two tidal humps running over th rotating Earth?

The lunar force of attraction, as you remembe causes the Earth's motion around the Earth Moon center of mass with the acceleratio $Gm_{☾}/a_{☾}^2$. However, it acts upon a body of mass located on the Earth nearer to the Moon wit a force $Gmm_{☾}/(a_{☾} - R_{\oplus})^2$ and on a similar bod located on the opposite side of the Earth, wit a force $Gmm_{☾}/(a_{☾} + R_{\oplus})^2$. The result is tha the first body should accelerate faster and th second body slower than the Earth does on th average. Therefore the tidal hump on the sid of the Earth turned to the Moon results from i own intensified attraction while that on the oppo site side stems from the fact that the Earth grav itation is stronger on the average than th gravitation of its farther side. These humps a sketched in Fig. 22.

The above discussion allows us to assess th distortion of the Earth's shape under the actio of the lunar attraction as follows: the heigh of the tide equals, by the order of magnitud

$$\delta R_{\oplus} \sim R_{\oplus}\frac{m_{☾}}{m_{\oplus}}\left(\frac{R_{\oplus}}{a_{☾}}\right)^3 \sim 0.36 \text{ m.}$$

Such estimate of the height of tides on the Eart is more or less true for a solid shell, the Earth' crust. Yet oceans, the density of which is lowe

than that of the Earth's crust, respond more actively to the tidal action than the land does. In high seas, however, the height of tides is merely about half a meter.

The observable height of tides increases significantly near coasts. The highest tides take place in contracting gulfs and shallow seas such as the Bering Sea and the Sea of Okhotsk

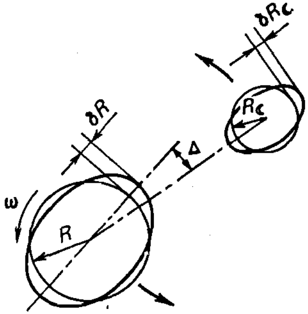


Fig. 22. Diagram of tides on a rotating planet and its satellite.

which absorb the energy of Pacific tides. In general, sea tides are especially high in latitudes $\varphi = 50°$ North and South. Tidal humps move there over the Earth's surface at a speed of $\omega_{\oplus}R_{\oplus}\cos\varphi \simeq 290$ m/s. This speed approximates the velocity of propagation of the longest oceanic waves. Thus, tides seem to be transported by oceanic waves and amplified on the way. The greatest difference between high and low water was registered in the Bay of Fundy, Canada, in latitude 45° North, where it reached 16.3 meters.

Sea tides are most apparent to us but one should not forget that the tidal action distorts the Earth's atmosphere as well: the atmosphere also slightly protracts towards the Moon and in the opposite direction. However, this atmospheric protraction causes alterations of atmospheric pressure near the surface of the Earth which are negligible compared to deviations from the average caused by weather changes.

Tides on the Earth are caused not only by the Moon but also by the Sun. The height of a solar tide can be estimated in the same way as that of the lunar:

$$\delta R_{\oplus}^{(\odot)} \sim R_{\oplus} \frac{m_{\odot}}{m_{\oplus}} \left(\frac{R_{\oplus}}{a_{\oplus}}\right)^3 \sim 0.16 \text{ m}.$$

It is thus clear that the Sun's influence on the tides is twice weaker than that of the Moon. The height of tides reaches its maximum when the solar and lunar actions overlap. It is clear that such a situation is possible at full moon and at new moon when the Sun, the Earth, and the Moon are positioned on one straight line.

The lunar attraction causes tides on the Earth. But does the Earth cause tides on the Moon? Estimate the distortion of the Moon's shape due to terrestrial gravitation:

$$\delta R_{☾} \sim R_{☾} \frac{m_{\oplus}}{m_{☾}} \left(\frac{R_{☾}}{a_{☾}}\right)^3 \sim 13 \text{ m}.$$

We shall return to the Moon's shape a little later: it is distorted much more than by a mere dozen meters.

There is one more and important effect of the close vicinity of a large satellite. No other body

in the solar system experiences presently such a strong influence as the Earth does: the Moon brakes the rotation of this planet!

Look again at Fig. 22. The angular velocity of the Earth's rotation is greater than that of the Moon's revolution around the planet. Therefore the Earth's rotation slightly drags the tidal humps. The humps appear not exactly on the straight line between the Earth and the Moon, they are slightly turned in the direction of the Earth's rotation. This angle of drag is assessed at 2°. This displacement brings about the asymmetry of the gravitation of tidal humps to the Moon which generates a minor moment of forces decelerating the Earth's rotation. Calculations indicate that the Earth's angular velocity should decrease by 2×10^{-10} of its magnitude per year. Consequently: the length of the day should increase each year by 2×10^{-5} s. Is there an experiment to register such a weak deceleration of rotation?

The relative accuracy of the modern atomic clock is 10^{-14}, i.e. its measuring error does not exceed 3×10^{-7} s per year. However, a direct measurement of the secular deceleration of the Earth's rotation is hindered by significant seasonal fluctuations of the velocity of rotation resulting mainly from atmospheric winds. The mass of the atmosphere accounts for 10^{-6} of the Earth's mass. Since atmosphere has no rigid connection with the solid part of the Earth, its rotation is partially independent. Therefore seasonal winds (monsoons) account for the Earth's sidereal day being by several ten-thousandths longer in April than in August. Thus, it is difficult to measure accurately a weak constant deceleration of rota-

tion against the background of such fluctuations

However, the first measurements of the secula deceleration of rotation were made when the atom ic clock was not yet available. The hand wa given by the Moon itself or rather by the sola eclipses which it causes. To understand the idea we have to consider the angle of the Earth's fal ling behind due to deceleration of rotation fo the period of, say, $t = 2000$ years $= 6.31 \times 10^{10}$ s. Let the value of deceleration be $-\dot{\omega} = 4.81 \times 10^{-22}\ s^{-2}$ which is the most accurat according to modern data. Similarly to the for mula of the path in uniformly retarded motion w get the value of the angle of lag $\delta = -\dot{\omega}t^2/2 =$ 55°. This is the angle of the Earth's turn i 3.5 hours.

If we assume now that the velocity of the Earth' rotation is constant and calculate back the char acteristics of the ancient total solar eclipses, w would arrive at the conclusion that those eclipse should have been observed by several dozens de grees farther westwards from the points wher they had been actually observed. It was that ver information on several solar eclipses which ha taken place in B.C. times which allowed to mak first estimates of the secular deceleration of th Earth's rotation.

The moment of forces decelerating the Earth originated by the Moon. At the same time increases the orbital moment of momentum of th Moon. Consider again Fig. 22: the tidal humps o the Earth drag the Moon. Due to this action th Moon slowly moves away from the Earth. Th average distance between them increases b

3 cm per year. It should be noted that the specific time of the evolution of the Earth-Moon system is of the same order of magnitude as the age of the solar system.

We do not know yet how the Moon was actually formed. However, it was certainly closer to the Earth in the past. Let us consider its shape from this point of view. The trajectories of the man-made satellites of the Moon have demonstrated that the center of mass of the Moon is shifted from its geometrical center towards the Earth by 2 or 3 kilometers and not by a dozen meters as required by the present equilibrium. This distortion of Moon's shape approached the equilibrium, when the Moon was positioned by 5 to 6 times closer to the Earth than it is today.

It is possible that billions of years ago powerful tidal forces had turned the Moon to face the Earth for ever by one side and Moon's body had "memorized" the great tide of that time. Note that the present shape of the Earth corresponds to its past velocity of rotation. But the Earth's "memory" is shorter: "merely" 10 million years.

The deceleration of the Earth's rotation alters its equilibrium shape from flattened at the poles to a more sphere-like one. This means that the mass should transit from equatorial areas to higher latitudes. However, the estimate of mass flow demonstrates that it is irrelevant compared to mass flow observed in the course of continental drift of the plates.

The kinetic energy of rotation decreases with the deceleration of the Earth's rotation. This results directly from the friction of tidal currents against the floor of seas and oceans. The shallower

the sea and the larger the surface over which th tidal wave was accelerating, the more intensiv the release of the tidal energy. The highes efficiency is observed in the shoals of the Berin Sea and the Sea of Okhotsk, the two of whic account for a quarter of the total power of tide of the World Ocean.

Let us estimate this power. The kinetic energ of the Earth's rotation equals

$$K = 0.33\,\frac{1}{2}\,m_{\oplus}R_{\oplus}^{2}\omega_{\oplus}^{2}$$

(the factor 0.33 is determined by the distributio of masses inside the planet: the density in th center is 2.4 times higher than the average densit of the Earth). The rate of decrease of this kineti energy, i.e.

$$|\dot{K}| = 0.33 m_{\oplus}R_{\oplus}^{2}\omega_{\oplus}\,|\dot{\omega}| = 2.8\times 10^{12}\ \mathrm{W},$$

is the power which the tides take away from th Earth's rotation. A part of this power equal t the ratio of the periods of the Earth's rotatio and the Moon's revolution (1/27.3 = 3.7 % i.e. 10^{11} W) is spent on the increase of the Moon' total energy in its motion away from the Earth The major part of the tidal power is released i the form of heat in bottom currents in shoal

The thermal power released by tides is presen ly insignificant compared to the energy whic the Earth receives from the Sun. But in the pas when the Moon was closer to the Earth and th planet rotated faster, tides could have bee making a significant contribution to the heat ba ance of the Earth's surface.

6. Why the Interior of Planets Is Hot?

It has been found out in the course of drilling boreholes and tunnels that the temperature of rocks increases with depth. This temperature rise is not constant and depends on the place where the holes are bored. The average increase approximates 30° per kilometer of immersion. Unfortunately, the accurate temperature data are available only for a thin layer of the Earth's surface since the deepest borehole has so far reached only 12 km, i.e. merely 0.2% of the Earth's radius. Nevertheless we can be sure that the temperature increase continues still lower. This is true because the heat flow, i.e. the power coming from the interior, almost does not change with depth. Let us estimate it.

The thermal conductivity of basalt can be employed for an assessment ($\varkappa = 2$ J/m·s·K). To find the heat flow over the entire Earth's surface ($4\pi R_{\oplus}^{2}$), it should be multiplied by the thermal conductivity $\varkappa$ and temperature gradient ($dT/dz = 0.03$ K/m). The heat flow will be

$$Q = 4\pi R_{\oplus}^{2}\varkappa\,\frac{dT}{dz} \simeq 3\times 10^{13}\ \mathrm{W}.$$

What is the cause and source of this flow of energy? One of the causes is energy release due to radioactivity. It is known that rocks contain a small but appreciable admixture of uranium. Its share is most significant in granites where it reaches several millionths of mass. The most common isotope of uranium is ^{238}U, which is the major contributor to the heat released by rocks. Each act of uranium decay liberates an α-particle

with the energy equal to 4.2 MeV. After flyin about 10 μm it stops and transfers all its energ to surrounding rock thereby heating it. Th half-life of ^{238}U equals 4.5×10^9 years. Hence i is easy to find out that energy release of pur uranium is 1.8×10^{-5} W/kg and granite i which uranium accounts for 10^{-6} in mass, release heat at the rate of 5×10^{-8} W/m^3.

A layer of granite about 20 km thick lies unde the continents. If there were no change in ura nium concentration with depth, it would suppl a significant share of the heat flow from the depths Hence comes a strong temptation to explain th total heat flow by the energy release assumin that uranium is present everywhere down to th depths of several thousand kilometers althoug its concentration is lower than in granite.

But most likely it is not so. Firstly, the hea flow under the bottom of the oceans turned out t be the same as in the continents, although there i no granite there, and the content of uranium i basalts is lower by the order of magnitude. Sec ondly, there is an amazing coincidence: all th elements which have long-living radioactiv isotopes (uranium, potassium, thorium, and stron tium) are in compounds (in the state of chemica equilibrium with rocks) easily soluble in water This means that in the course of rock evolutio they were migrating with water and now thei concentration can be significant only in rock containing free or crystal water. This limits th depth of propagation of long-living radioactiv isotopes by the same 15 to 20 km.

Thirdly, the principal reason for the unsound ness of the radioactive hypothesis consists in tha

there is another powerful source of the thermal energy of planet interiors: the gravitational. This source stems from the times in the very beginning of the history of the Earth and planets when they were only forming. The formation of planets was accompanied by collisions, mutual braking, and amalgamations of minor celestial bodies. For brevity all such celestial bodies, from cosmic specks of dust to minor planets, will be referred to as asteroids. The scars left from the times when the planets were forming can be seen even by a naked eye on the Moon's surface. These are lunar seas: circular lowlands with a diameter reaching a quarter of the Moon's disk and filled by dark basalt lavas. The seas are in fact the marks left by impacts of large asteroids with dimensions of dozens kilometers with which the Moon collided about 4 billion years ago.

The lunar seas are underlied by high-density layers resulting from the shock compression of rocks by falling asteroids. These high-density layers are called "mascons" (the term is an abbreviation of "massive concentration"). The largest of mascons reach by mass 10^{-5} of the Moon's mass being positioned at the depth of 50 km under its surface. These very mascons, by the way, make it impossible to launch a satellite with a service life longer than five years into the Moon's orbit since satellite's stability is adversely affected by the gravitational field of the Earth and low orbits are unstable because of mascons: their perilune gradually slides down to a tangency with the Moon's surface.

A process to some extent similar to that gravitational retardation of the satellite could prob-

ably have taken place at the initial and principal stage of the formation of planets and the Sun itself. There was a swarm of celestial bodies, a system of gravitationally connected asteroids, all of which were like satellites of one another. The gravitational field of such a system is inhomogeneous with every body moving in chaotic potential and gradually spending energy on multiple tidal interactions. The swarm was progressively compressing and catching by the same gravitational retardation foreign asteroids, which were flying through and by the system. The captives became member-satellites thereby increasing the mass of the swarm. Unfortunately, all these processes are so complicated and versatile that we have so far failed to calculate the entire process of the formation of planets from the start to the end.

Such an amalgamation of a swarm of bodies into a planet cannot be strictly separated from the increase in its mass due to direct collisions with asteroids. The amalgamation of gravitationally connected bodies is finished by a bombardment of a young planet's surface. This very bombardment left its traces on the surfaces of the Moon, Mercury, and Mars.

As regards the Earth, there are no mascons and craters, as old as those of the Moon, on this planet. The Earth's surface is continuously renovated by the drift of continental plates and the active atmosphere and oceans do not take much time to wash away or erode craters. The possibility to locate about a hundred of heavily time-smoothed circular structures—up to a hundred kilometers in diameter—was given only recently by contrast photos taken from space.

With a lesser confidence can we claim a connection between asteroid impacts and such circular structures as a formation 240 kilometers in diameter and 800 meters deep in Antarctica, in Wilkes Land, and a formation 440 km in diameter on the eastern coast of Hudson Bay, the half of which can be seen in geographical maps. The greatest easily detectable crater is located in Arizona, USA. It is 1265 m in diameter and 175 m deep and was formed merely 25 to 30 thousand years ago as a result of the fall of a body with a mass of approximately 2×10^9 kg.

The velocities with which celestial bodies fall to the Earth are always higher than the escape velocity (i.e. 11 km/s) and approach, generally, 20 km/s which is little less than the orbital velocity of the Earth. This is due to the fact that the majority of asteroids, which have a chance of colliding with the Earth, are moving in the same direction and by an orbit more elongated than that of the Earth. It is a rare thing that the velocity of a body entering the Earth's atmosphere may reach 53 km/s which is possible only in the case of collisions with comets. If the velocity is that high, the nuclei of comets do not reach the Earth's surface and burn completely in the atmosphere.

In the collision of an asteroid with a planet the major portion of its kinetic energy is spent on the retardation in the atmosphere, heating of rocks in the impact point, and ejection of rocks from the crater. The mass of a material ejected from the crater exceeds the mass of the asteroid by a factor of one to three hundred and the velocities of individual drops of melted rocks thrown out of

the crater may reach several kilometers per second. We can estimate the magnitude of this velocity with the help of several meteorites which were rather unexpectedly found on the Earth and infallibly identified as lunar rocks. Their lunar origin means that they were ejected with a velocity higher than 2.4 km/s, the lunar escape velocity, from a crater formed on the Moon, escaped it, and became the Earth's satellites to fall later—probably after a long period of time—to the Earth.

The total irrevocable loss of mass due to the splitting of asteroids into fine particles can be extremely great, up to a total disintegration of colliding bodies if their masses are close and the velocity of collision is high. The shell and internal parts of large asteroids may differ significantly by chemical composition. Therefore catastrophic collisions of asteroids may to some extent account for the diversity of the chemical composition and crystalline structure of meteorites which get to the Earth.

The rare celestial visitors, C-type carbonaceous meteorites, are considered the initial material of the solar system which has not been transformed in collisions. Their chemical composition is similar in proportions to that of the Sun excluding volatile noble gases and hydrogen. Among other things they contain up to 20% of water in crystalline form, several per cent of carbon and sulphur. A slight heating removes water from the C-type meteorites and leaves a material containing 34% of ferric oxide (FeO), 33% of silicon oxide (SiO_2), 23% of magnesium oxide (MgO). These very elements but in different proportions deter-

mine the composition of other meteorites, asteroids, and planets of the terrestrial group.

The most common meteorites are stone meteorites, the so-called chondrites. Their silicate structure contains chondri, the grains with the diameter of several millimeters. We have failed so far to achieve such a crystallization in laboratory conditions on Earth. The chondrites are sometimes impregnated with the alloy of iron and nickel. Their share varies, but the contents of iron is usually less than in heated C-type meteorites.

Finally, the third type of meteorites are stony-iron meteorites. The major component of these meteorites, in contrast to the above-mentioned, is an iron-nickel alloy and stone impregnations are just a hardened melt without a chondral structure.

The most natural (although exposed to objections) hypothesis is that iron and stony-iron meteorites were born in cataclysmic collisions of asteroids, i.e. formed from their central parts which had enriched them with heavy iron; stone meteorites are splits of the upper layers of asteroids and minor planets chopped off in weak collisions.

In the collisions of bodies significantly different in mass—such as a small asteroid and an embryo of a planet—the mass of lost splits is not great and the kinetic energy of collision is released mainly near the larger body. The material ejected by the impact is dispersed in the vicinity of the crater heating the surface of the planet by secondary hits. This energy is further transported—at rather high rate—back to the space by the thermal radiation. The formation of

mascons, the packings at significant depths, is caused only by falls of large asteroids but even in that case the portion of energy released deep into interior is insignificant compared to the energy heating the surface of the prototype planet.

At first sight this should mean that when a planet is being formed of asteroids its interior may remain cold. It may seem to suffice therefore that the surface could cool down in the time interval between the formation of neighbouring craters. But this is so only at first sight.

Each collision increases the mass of a planet's embryo. As soon as it would exceed the critical mass, more than 10^{20} kg (cf. Table 4), the prototype planet would start to turn spheric under the action of gravitation. The material near the center of the sphere is being inelastically condensed while the gravitational energy of the entire planet decreases and the excessive energy is released in the form of heat directly in the interiors. The heating induces a partial melting and starts chemical reactions. The heavy iron-containing minerals submerge in the melt towards the center while the lighter silicate components are displaced upwards to the mantle-shell. The gases, both produced by heating and inert, previously impregnated in minerals, are intensively released. This gravitational differentiation of the material liberates thermal energy even greater than the energy of simple compression. If the mass of the protoplanet reaches approximately 10^{28} kg, the liberated energy suffices to melt or soften the majority of minerals contained in carbonaceous chondrites.

If the final mass distribution of a planet is known, the total gravitational energy released when the planet was formed of minor bodies, the remote asteroids, can be calculated. The gravitational energy is negative. Therefore it is logically easier to calculate the energy required for a disintegration of the ready planet into small parts. A breaking away of a minor mass Δm from a planet with mass m and radius R requires the energy equal to $Gm\Delta m/R$. This very energy should be calculated by gradually decreasing the values of mass and radius.

The present distribution of masses inside the Earth is known sufficiently well from seismic data, the times of sound propagation by various trajectories inside the Earth. There is a solid sphere at the planet's center with the radius of 1217 km and density of about 13 g/cm^3. Further upwards the planet's material is melted up to the radius of 3486 km. To this fact testifies the observation that only longitudinal sound waves pass through this area and the transverse waves fail to propagate. The density of this melt falls gradually from 12.1 to 9.9 g/cm^3. The material becomes solid at the outer boundary of the liquid nucleus and density falls down abruptly to 5.5 g/cm^3. The density decreases further evenly or by small jumps to reach the average 2.7 g/cm^3 on the Earth's surface. If we assume that the central solid nucleus is composed of iron and the melt is made up by iron oxide FeO and iron sulfide FeS, the Earth's chemical composition will be similar on the average to that of carbonaceous chondrites.

The calculation of the Earth's total gravita-

tional energy on the basis of this distribution yields

$$E_{\oplus}=0.388\,\frac{Gm_{\oplus}^{2}}{R_{\oplus}}=-1.45\times10^{32}\ \mathrm{J}.$$

The same amount of thermal energy, but with the positive sign, was released in the course of the Earth's formation.

However, as it was mentioned above, not all that energy is liberated in the depths. A considerable portion of it remains on the protoplanet's surface after impacts of asteroids and is irradiated into space by thermal radiation. To find this radiated share of the Earth's total gravitational energy, calculate the gravitational energy of an imaginary planet with equal mass but composed of incompressible material. We assume that the total gravitational energy liberated during the formation of this imaginary planet can be irradiated into space and the interior would remain cold. The density of the imaginary planet does not depend on the radius since its material is incompressible. So let the density of this planet be equal to that of carbonaceous chondrites: $\rho_* = 3\ \mathrm{g/cm^3}$. In that case the radius of the imaginary Earth would be $R_* = (3m_{\oplus}/4\pi\rho_*)^{1/3} =$ 7800 km and its gravitational energy amounts to

$$E_*=-\frac{3}{10}\,\frac{Gm_{\oplus}^{2}}{R_*}=9.2\times10^{31}\ \mathrm{J}.$$

This value should be subtracted from the total gravitational energy of the Earth $E_{\oplus}$. The result would be that it had taken 5.3×10^{31} J to heat the Earth's interior.

Is it much or little? To answer this question we have to decide whether this energy would suffice to melt the Earth's nucleus consisting, assumingly, of ferric oxide. An assessment of the energy required for melting is no easy task since melting points and specific energies of melting of all substances rise with pressure. Nevertheless a rough estimate of the calculated value is approximately twice higher than the present store of the thermal energy contained in the melted and heated to 3000 to 4000 K nucleus of the Earth and its colder solid mantle.

The thermal energy of 5×10^{31} J had certainly been liberated inside the Earth not momentarily. The gravitational differentiation moves masses of different chemical composition to distances comparable to the Earth's radius and this takes time. The differentiation is slowly proceeding even nowadays, but the most intensive division of the Earth into iron nucleus and silicate mantle was taking place soon after the planet's formation, in the first billion years of its existence.

The gravitational differentiation both liberates the energy in the interior and gives hand to deliver it to the surface: light fractions rise upwards and carry heat from the depths. This transport of heat and masses resulted in continuous eruptions of a multitude of volcanoes on the surface. Ejected gases had formed the Earth's initial atmosphere and lavas had shaped the basalt crust. Because of those eruptions heat losses from the depths were much higher in the first billion years than they are now.

We have estimated the present heat flow by heat conduction of rocks and temperature differ-

ences detected in them. This, however, is not a comprehensive assessment. In addition to the heat conduction there is a convective heat exchange when heated rocks rise at the boundaries of the plates where the new crust is being formed and when the cooled crust submerges at the opposite boundaries of plates. To measure this heat flow would be quite a difficult operation. Assumingly, it is now of the same order of magnitude as the heat conduction flow which amounts to 3×10^{13} for the entire Earth.

Rough calculations indicate that about a half of the total energy liberated and converted into heat of the gravitational energy was transported to the surface of the Earth already in the first billion years of its existence. Hence the average heat loss of the Earth's interior can be assessed at approximately 10^{15} W. We shall see further that such a heat flow would suffice to solve the Earth's climate paradox in its first billion years, i.e. the paradox effectuated by the reduced luminous emittance of the Sun at that time.

So let us turn to the Sun, the principal source of the energy heating the Earth.

Chapter 3

The Sun: a Source of Energy

1. First Information on the Sun

Further description of the Earth, its atmosphere, oceans, and climate is impossible without a reference to the principal source of energy of the solar system and the major star, the Sun. Let us make a brief armchair tour to the Sun, look into its interior, where the solar energy is born by nuclear reactions. A certain preparation should precede any travel, even a theoretical one. For this travel we shall need some luggage consisting of preliminary knowledge of the Sun, the information which you may be well familiar with.

As you know, the Sun is removed from the Earth to the average distance $a_{\oplus} = 1.496 \times 10^{11}$ m. From that distance it looks like a bright disk with angular diameter slightly more than half a degree, more precisely, 9.3×10^{-3} radians. This being known, the Sun's radius is easily calculable: $R_{\odot} = 6.96 \times 10^{8}$ m which is 109 times more than that of the Earth.

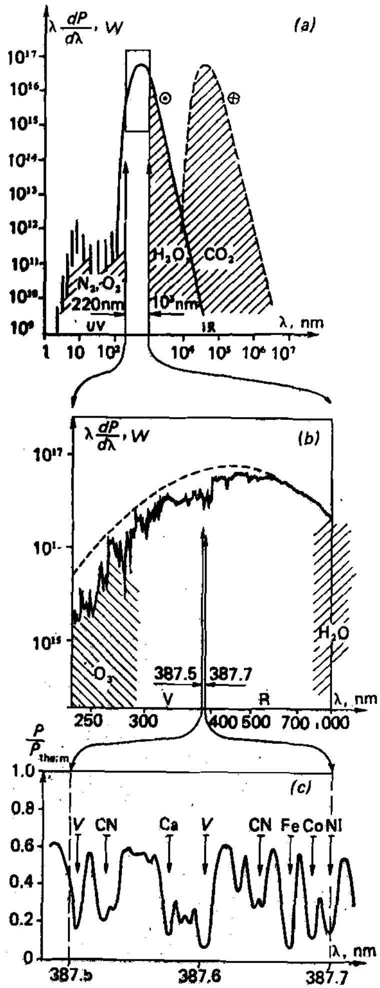
We know the mass of the Sun from the law of gravitation and motion of the planets: $m_{\odot} = 1.99 \times 10^{30}$ kg. This very law can be employed to find the free fall acceleration on the solar surface: $g_{\odot} = Gm_{\odot}/R_{\odot}^2 = 274$ m/s^2, which is almost thirty times more than the free fall acceleration on the Earth. Let us calculate now the

mean density of our star:

$$\rho_\odot = \frac{3m_\odot}{4\pi R_\odot^3} = 1.41 \times 10^3 \text{ kg/m}^3.$$

At first sight this result may seem surprising The mass of the Sun and free fall acceleration o its surface being so great, the solar materia should be highly compressed, but in fact it mean density is slightly above that of wate and notably lower than the mean density of th Earth equal to $\rho_\oplus = 5.52 \times 10^3$ kg/m³. W shall see that the reason therefore is the lo molecular mass of the solar gas and the hig temperature of solar interior. What can be sai about the temperature of the Sun from thi cursory study?

Recall the laws of the thermal radiation whic were treated in Chapter 1. To find the temperatur on the surface of the Sun one should first of al examine the spectrum of the solar light and determine what share of energy falls on the mino range of frequencies $d\omega$ for all the frequencies o the solar radiation. Consider Fig. 23*a*. The sola spectrum turns out to be rather close to Planck's spectrum of equilibrium thermal radiation and the temperature of the Sun's surface is $T_\odot =$ 5780 K. Interestingly, at such a temperatur the maximum energy is irradiated in that sectio of the spectrum to which the human eye is most sensitive: in the wavelength range between 40 and 800 nm. Is it a coincidence? Hardly so. The eyes of the living beings should probably adap to the brightest section of the spectrum of the light they live in.



Another experiment can also be employed find the temperature of the Sun's surface. W could measure the energy delivered in the enti spectrum by the solar rays to a unit area p second, say, one square meter of the Earth surface facing the Sun. This value, $s_{\odot}$, is ve important, as we shall see, for the explanatio of the Earth's climate. It turns out to be pract cally constant: the energy flux from the Sun do not change. Hence the value $s_{\odot}$ is called th solar constant. It is better measured from a sate lite outside the Earth's atmosphere which absor a part of the solar radiation. The solar consta equals 1.36×10^3 W/m^2.

Such flux of energy falls on the Earth's surfa facing the Sun. But the Sun sends equal quantit

Fig. 23 (*a*). Spectral distributions of the solar radiatio coming to the Earth and the Earth's radiation. Gener view of spectra near their maxima approaches the the mal spectra: 5870 K for the Sun, 258 K for the Earth In the ultraviolet the Sun's spectrum is nonthermal an changes significantly with time: it is indicated by vert cal hatches. The ultraviolet part of the solar spectru is absorbed by nitrogen and oxygen of the Earth's atm sphere. The spectrum of the Earth's radiation lies i the infrared zone in which light is absorbed by wat vapour and carbon dioxide. (*b*) An area of the sol spectrum near the visible spectrum. The red zone o the spectrum approaches the thermal one (dashed line the violet zone differs notably from them. At 220 < $\lambda < 290$ nm the solar radiation is absorbed by th ozone of the Earth's atmosphere. (*c*) A complicate microstructure of the solar spectrum. The absorptio of light by atoms and molecules in the Sun's atmosphe forms in the visible spectrum more than 20 thousan spectral minima—the dark Fraunhofer lines. They ar shown in detail in a spectral area within the range o merely 0.2 nm.

of light in all directions and, consequently, such is the power of solar rays passing through each square meter of the entire sphere with radius $a_{\oplus}$ around the Sun. To find the energy irradiated by the Sun per second, $s_{\odot}$ should be multiplied by the surface area of this sphere. This value is called solar luminosity: $L_{\odot} = 4\pi a_{\oplus}^2 s_{\odot} = 3.83 \times 10^{26}$ W.

The entire flux of solar energy passes, certainly, through the solar surface as well. What is the energy flux density there, i.e. what power is emitted by a square meter of the Sun's surface? Let us divide the solar luminosity by the area of the Sun's surface $4\pi R_{\odot}^2$. The result will be that the radiation flux density, or, in other words, the Sun's brightness, equals $S = L_{\odot}/4\pi R_{\odot}^2 = 6.29 \times 10^7$ W/m^2.

But we know the Boltzmann principle relating the flux density S to the temperature of the emitting surface. Therefore, we can employ another method to calculate the Sun's surface temperature. The result equals that established with the help of the solar spectrum: $T_{\odot} = (S/\sigma)^{1/4} = 5780$ K.

The solar spectrum, however, differs from the blackbody radiation spectrum even in that area of Fig. 23*a* where they are shown to coincide. This difference carries the information on the presence of chemical elements on the Sun. The matter is that each atom has its specific and personal spectrum of radiation composed of very narrow lines. As early as at the beginning of the last century a German optician J. Fraunhofer discovered a multitude of dark lines in the solar spectrum. It turned out therewith that the

wavelengths at which the dark Fraunhofer line were detected in the spectrum accurately corre pond to the spectra of chemical elements. Th reason why the Fraunhofer lines are dark will b treated in the section discussing the surface o the Sun. For the time being the important thin is that different intensities of the lines of variou elements have allowed to establish the relativ contents of the respective elements near the sola surface. Considering that both the atmospher and interior of the Sun are highly mingled dow to great depths, this very composition can b regarded approaching the average compositio of the Sun.

The Sun has turned out to be composed, by number of atoms, of hydrogen (app. 91%) an helium (app. 9%), while the presence of othe elements is insignificant. The helium's shar by mass is notably greater: 27% since the mas of a helium atom exceeds that of a hydroge atom by a factor of four. Interestingly, heliu was first discovered on the Sun exactly due t its Fraunhofer spectrum, which did not matc any other element. On the Earth helium wa discovered much later. The name of this elemen stems from the Greek *helios*, sun.

Recall now that approximately the same mas portion of helium was predicted by the cosmolog theory of the element formation treated in Cha ter 1. But the initial, prestar matter compositio was devoid of any heavy elements while almo all the elements of the periodic table are prese on the Sun together with hydrogen and heliu Their total mass, however, is insignificant, le than 3%. The major "share-holders" among tho

elements are oxygen, nitrogen, and carbon. Consider Fig. 24, demonstrating the distribution of elements over the entire solar system, bearing in mind that the Sun accounts for the overwhelming part of its mass. It is interesting that the

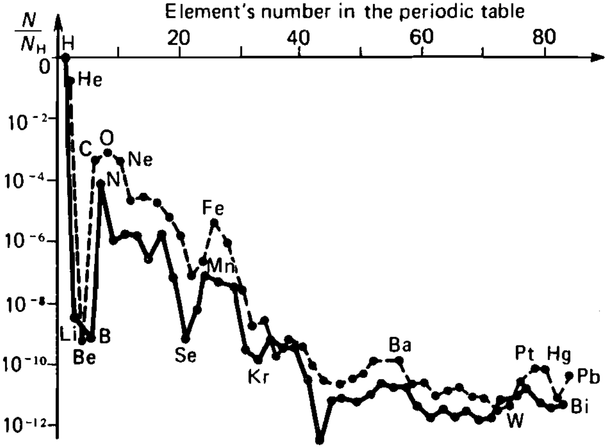


Fig. 24. Occurrence of elements in the solar system according to the occurrence of hydrogen. Continuous lines connect elements with odd numbers, dashed lines connect those with even numbers.

number of elements with even numbers is ten times that of their table-neighbours with odd numbers. The qualitative nature of this dependence will become clear when we shall discuss the mechanism of the formation of other elements from hydrogen and helium which takes place inside the stars.

2. The Sun's Interior

Why do the stars shine? Where does the immense energy, which the solar surface continuously emits, come from? What is the internal structure of the Sun? These questions may seem senseless and unanswerable since no man can visit the Sun. Even man-made instruments cannot endure the enormous temperatures and pressures of the solar interior. Nevertheless, quite adequate information is already available on many processes taking place inside the stars and the Sun. Strange it may seem, but the fact is easily comprehensible.

The possibility to study the interior of the stars is provided by theory. But can one rely upon theoretical speculations supported by the experience of the residents of the Earth? The conditions in which experiments are arranged on the Earth are much different from those of real stars. The answer is that we should try to calculate the structure of the Sun assuming that our theoretical conjectures are true. If the results would be reasonable and agree with observations, it would mean that major theoretical premises are also true in unusual conditions. Let us try and see.

We shall begin with an attempt to find a dependence between pressure, temperature, and density of the material the Sun is composed of. The gas on the solar surface consists mainly of neutral atoms since hydrogen molecules disintegrate into individual atoms at the temperature as high as that on the surface of the Sun. Soon we shall see that the temperature inside the

Sun is enormous not only by terrestrial standards, but also compared to the temperature of its surface. Even at small depth the temperature is so high that not only molecules but also atoms disintegrate. Electrons of atom shells break away from nuclei and gas turns into plasma.

Plasma is an electrically neutral mixture of negatively charged electrons and positively charged nuclei of atoms. Experiments carried out on the Earth indicate that plasma can be reliably considered as a mixture of ideal gases: the gas of electrons and the gas of nuclei. For the sake of simplicity we shall assume that the solar plasma consists only of three types of particles: there are 9 helium nuclei per each 91 protons (hydrogen nuclei) and 109 electrons broken away from the nuclei.

You know the Mendeleev-Clapeyron equation relating pressure, density, and temperature of an ideal gas. Write it down not for the density ρ but for the concentration of particles n. The usual density of each individual gas is a product of the concentration of its particles by the mass of one particle m:

$$p = nkT; \quad \rho = nm.$$

One of the right-hand-side multipliers is Boltzmann constant $k = 1.38 \times 10^{-23}$ J/K.

These equations are true for each ideal gas of the mixture. We do not need an equation relating overall pressure to overall density and temperature. The temperature of all gases in equilibrium is the same. The overall density is the sum of individual densities. The overall pressure is the

sum of individual or, in other words, partial pressures.

The mass of helium nucleus exceeds that of proton by a factor of four and electron's mass is negligible compared to them. Finally, we know that the densities of particles (electrons (n_e), protons (n_p), and helium nuclei (n_α)) are related to one another as $n_e : n_p : n_\alpha = (2 - X) : X : (1 - X) = 109 : 91 : 9$. The X here indicates the share of protons. Let us tabulate now what has been said (cf. Table 6).

Considering this table, it is useful to recall that the Avogadro constant (6.03×10^{26} kg^{-1})

Table 6

Equations of State for the Gases in the Sun's Plasma

	Concentration of particles	Pressure	Density
Electrons	$n_e = \frac{2-X}{3-X} n$	$p_e = \frac{2-X}{3-X} nkT$	$\rho_e = 0$
Protons	$n_p = \frac{X}{3-X} n$	$p_p = \frac{X}{3-X} nkT$	$\rho_p = m_p \frac{X}{3-X} n$
Helium nuclei	$n_\alpha = \frac{1-X}{3-X} n$	$p_\alpha = \frac{1-X}{3-X} nkT$	$\rho_\alpha = 4m_p \frac{1-X}{3-X} n$
Sum	$n = n_e + n_p + n_\alpha$	$p = nkT$	$\rho = \frac{4-3X}{3-X} m_p n$
Result	$p = nkT = \frac{\rho}{\mu m_p} kT$; $\mu = \frac{4-3X}{3-X} = 0.61$		

is practically equal to m_p^{-1} and its product by the Boltzmann constant is the gas constant equal to 8.31 J/(kg·K). You would see then that the equation which we have got for the solar material is a usual Mendeleev-Clapeyron equation for a gas with molecular mass $\mu = 0.61$.

The resulting small value of the average molecular mass (less than one) stems from the small mass of the electron. The electron component of plasma does not make a significant contribution to the density but produces pressure higher than that of protons and helium taken together. The low molecular mass of the solar gas is one of the reasons for the mean density of the Sun being that low.

Thus we have got an equation describing the plasma of the deep solar interior. We know the mean density of the Sun. Let us try to estimate the pressure and temperature in the central areas of the Sun by the order of magnitude. We have estimated the pressure in the interior of the massive celestial bodies in the previous chapter. The same considerations are applicable to the Sun. The material inside it is compressed by the gravitational attraction. Therefore we can employ the formulas of the pressure inside the planets for a rough estimate of the mean pressure inside the Sun:

$$p_\odot \sim \rho_\odot g_\odot R_\odot \sim \frac{G m_\odot^2}{R_\odot^4} \sim 10^{15}\ \mathrm{N/m^2}.$$

Knowing the mean pressure and mean density and using the equation of the solar plasma we can easily estimate the temperature of the central

areas of the Sun:

$$T_S \sim \frac{p_\odot m_p}{k\rho_\odot} \sim \frac{G m_\odot m_p}{k R_\odot} \sim 2 \times 10^7 \text{ K}.$$

Twenty million kelvins. This is certainly only an estimate but it does not differ very much from the results of accurate calculations. But what techniques are employed to calculate the internal structure of the Sun and other stars?

Accurate calculations treat the entire volume of the star as composed of a multitude of thin spheric shells. Each of the layers has its special pressure, temperature, and density *slightly* different from the identical values in the adjacent layers. These minor alterations are further calculated from layer to layer with regard to nuclear reactions releasing energy, transfer of this energy from depths to the surface, gravitational attraction to the center, and the pressure of outer shells. These equations of minor increments in all relevant values from layer to layer are termed differential equations.

The equations being compiled, the problem turns purely mathematical and can be solved on a computer, the dependences of all the values on the radius should be calculated including temperature $T(r)$, density $\rho(r)$, pressure $p(r)$, gravitational acceleration $g(r)$, and mass $m(r)$ within the limits of radius r. Figure 25 gives an example of the results of such a calculation.

Note an interesting peculiarity: the gravitational acceleration reaches its maximum not on the surface but at a rather significant depth, at the

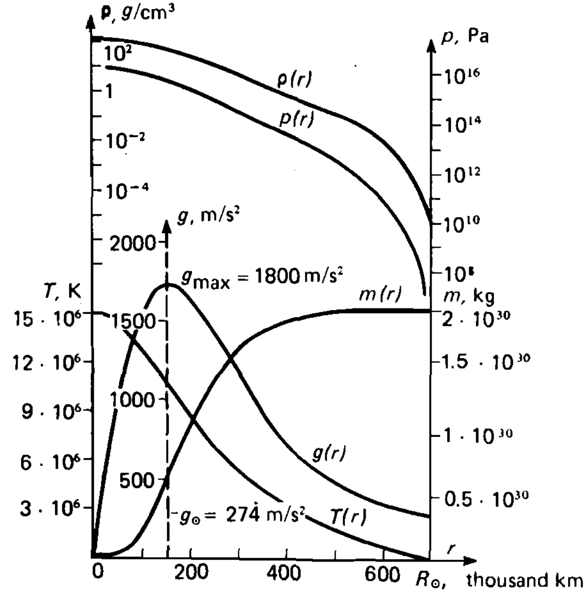


Fig. 25. Alteration of temperature T, acceleration g, density ρ, pressure p and mass contained within radius r inside the Sun.

radius equal to merely 0.217 radius of the Sun*. The maximum acceleration is 6.5 times higher than the acceleration on the surface of the Sun, $g_\odot$. The density, pressure, and temperature

* This statement holds true not only for the Sun but also for the majority of planets. On the Earth the gravitational acceleration reaches its maximum not on the surface but at the depth of 2900 km, where it equals 10.69 m/s² which is by 9% more than that on the surface.

decrease monotonically with radius. However, the pressure and density fall rapidly near the very surface of the Sun while the temperature starts to decrease very rapidly already at the distance of 1/5 solar radius from the center. This makes it possible to point out the central area of the Sun which is called solar core or the nucleus. The nucleus is the hottest part of the Sun with almost homogeneous high density exceeding the mean value by a factor of hundred. This very nucleus is the source of almost all solar energy and the remaining part is the "blanket" of the nucleus keeping it from cooling down and slowly conducting the energy of the central area up to the surface.

Are these calculations accurate enough, namely is our knowledge of solar depths adequate? The major difficulty in the construction of the models of the Sun *is that the comprehensive information* on the chemical composition of the interior is unavailable: the contents of heavy elements in the nucleus of the Sun may be higher than on its surface. This composition is also calculated proceeding from various assumptions on the mixing of the solar plasma but uncertainty remains yet high. Nevertheless, it is believed that the accuracy of our knowledge about the solar depths approaches 10%. This assumption *is supported by the fact that the* temperature and density of the material in the nucleus of the Sun have turned out exactly such as required to support nuclear reactions there.

The temperature is so high deep inside the Sun that all material is totally ionized and consists of atomic nuclei and free electrons. However,

the rise to the solar surface is accompanied by two successive recombinations of helium and one recombination of hydrogen. In the process of recombination electrically neutral atoms form from ions and electrons.

The mean molecular mass μ increases with recombination. If $\mu < 1$ in the Sun's core due to the contribution of light electrons, it equals $\mu = 3 - 2X = 1.27$ *on the surface where free electrons* are scarce. In other equal conditions the gas with a greater molecular mass has a higher density. This situation favours convection in the Sun's shell at the distance from the center $r > 0.7R_{\odot}$.

What ~~is convection~~vection? Let us leave the Sun for a moment and turn to our everyday experience. If you look into the distance above a fire you would notice that the shapes of various things fade out and tremble. You have seen this happen above hot asphalt on a highway at summer midday, and above a pan with water placed on a stove. Twinkling of stars is also an effect of convection in the Earth's atmosphere: warm jets rise and cold ones descend to take their place. We can see this motion because the path of a ray in such a mingling medium is not straight: it curves when passing the boundaries of jets.

Conditions required for convection to occur are far from being apparent. It turns out that the condition of the upper medium being colder than the lower one does not suffice. Convection requires that the temperature drop with altitude exceeded a certain limit. The increase in the mean molecular mass with height contributes much to convection since a potentially heavier gas takes the upper position. On the Sun it hap-

pens in recombination and in the Earth's atmo-sphere it occurs when water vapour condensate in clouds.

3. The Energy of Nuclear Reactions

At high temperatures and pressures the nucle of light elements may join one another in colli-sions thereby forming the nuclei of new heavie elements.

The source of energy of the stars was unknow in the last century. In 1905 A. Einstein produce his theory of relativity and discovered that th principle of conservation of energy and the prin-ciple of conservation of mass, well known fro chemistry, are in fact a unified law. The mass o a system may turn out different before and afte the interaction but the difference of masses, o in other words, the mass defect Δm is fully com-pensated by the change in the system's kineti energy: $\Delta E = -mc^2$. Thus the sum $E + mc^2$ the total energy of the system, is accuratel conserved in any reactions.

There was no way to test this statement b experiments until E. Rutherford discovered th possibility of nuclear conversions, i.e. nuclea reactions, in 1918. And that is what the Germa physicist W. Nernst wrote as early as 192 "The relationship established by Einstein pro-vides a basis for further conclusions of muc greater importance. From this point of view th radioactive disintegration is one of the possibili-ties to extract enormous quantities of energy fro matter. A technical exploitation of this store energy does not seem principally impracticabl

Just recently Rutherford has apparently produced that type of energy, however on a microscopic scale, when he managed to decompose nitrogen by way of radioactive splitting. Yet no illusion should be made that the technical generation of the above-mentioned energy is a matter of the immediate future which would depreciate coal. On the other hand, there is no objection to one of the most serious technical problems being developed thereby". How can one fail to be im-pressed by the progress of science in the last 80 years and the deep insight which some scien-tists have demonstrated at the very start of that way!

The principle of conservation of energy and mass holds true for chemical reactions as well. In those reactions, however, the change in the mass of reaction products compared to the initial mass is so small that it is almost undetect-able by experiments.

Let us actually consider the decrease in mass in the course of, say, chemical reaction of coal burning. The specific heat of reaction

$$C + O_2 \rightarrow CO_2$$

is $q = 3.3 \times 10^7$ J/kg. This means that the burning of 12 kg of coal, with which 32 kg of oxygen react, liberates 3.96×10^8 J of energy. It is equivalent to 4.4×10^{-9} kg of mass. This is exactly the figure by which the mass of carbon and oxygen prior to reaction exceeds the mass of the produced carbon dioxide. But is there a way to detect a relative decrease that small (10^{-10}) in the mass of reagents? There is none, but nobody ever tried to do that because of its apparent use-

lessness. There is a sufficient proof of the pri ciple of equivalence between mass and energ in nuclear physics. The principle of mass co servation remains to hold true in chemistry well as in everyday life: the mass of matter conserved with the enormous accuracy of 10^{-1} As regards the nuclear reactions, the relati change of mass in the transformations of elemen is quite observable. Therefore the energy relea in nuclear reactions is much higher than that chemical reactions.

The reader certainly knows that atomic nucl can be considered consisting of positively charge protons and electrically neutral neutrons. Th number of protons in the nucleus equals th ordinal number of an element. A neutral ato houses equal number of negatively charge electrons and protons. In atoms electrons mo around nuclei at the distance between 10^{-10} an 10^{-9} m. They determine chemical properties elements. A change in the number of neutro in the nucleus does not affect chemical natu of an atom. It produces nuclear modifications elements, isotopes which differ from the pare element by the nucleus mass.

As you know, the atomic masses of the majorit of chemical elements approach integer number The atomic masses of some isotopes are ev closer to integer numbers. This is so because t proton mass m_p approaches the mass of neutr m_n and the binding energy of these partic inside the nucleus is less than $m_p c^2$.

The binding energy is generated by nucle forces. These forces are the forces of attracti acting between the particles of atomic nucl

protons, and neutrons. In this case the nuclear attraction almost does not depend upon the type of particle: the interaction of two neutrons is similar to that of two protons and coincides with the nuclear interaction between a proton and a neutron. But in contrast to electromagnetic forces the nuclear forces act only at very short distances.

The range of nuclear interaction $r_0 \simeq 1.5 \times 10^{-15}$ m approximately coincides with dimensions of the proton and the neutron. If several protons and neutrons are removed to distances of about r_0, they are unified by nuclear forces into a compact group: the atomic nucleus. The volume of atomic nuclei increases proportionally to the total number of protons and neutrons in the nucleus. This property of atomic nuclei reminds of the drops of incompressible liquid the volume of which is also proportional to mass. The dimensions of nuclei are thousands times less than the dimensions of atoms. Even the size of heavy nuclei is about several radii of nuclear interaction r_0.

The attraction of nuclear forces by far exceeds the electric repulsion of protons at distances that short. But with distances reaching several nucleus diameters nuclear forces become extremely weak. When atomic nuclei are rearranged in the course of nuclear reactions, the short-range but powerful forces of attraction produce a rather significant work at the distance of about r_0. This very work originates the mass defect. Simultaneously, the decrease in total mass in the course of nuclear reactions results in the liberation of energy. That is the kinetic energy of reaction products, the newly formed nuclei, and the energy

of particles, in particular, the energy of photon a quantum of electromagnetic radiation release by the reaction. In the final analysis, the majo portion of all energy liberated by nuclear reac tions, which take place inside the stars, is con verted into thermal energy. The fast-movin particles are retarded by the surrounding matter photons are also absorbed by plasma.

A reaction equation should conform to severa principles of conservation. Firstly, these includ the principle of conservation of electric charge the algebraic sum of charges should be identica in the left-hand and right-hand parts of th equation. Secondly, the total number of free an nucleus-bound protons and neutrons should re main constant. The latter is called the baryo number. Finally, the last rule reads: if an elec tron (e) or its antiparticle, positron (e^+), appea in the course of a reaction, a neutrino shoul also emerge. Electrons, positrons, and neutrino have a collective name of leptons and the prin ciple just mentioned is called the conservation o lepton charge.

Neutrino is a particle with a zero electric charg and a zero mass. It practically does not interac with other particles and flies away at the velocit of light immediately after its birth. Neutrino formed inside the Sun penetrate its entire bod without a delay, dispersion, and absorption. Th energy which they carry away is never trans formed into heat. However, we are most intereste in that part of nuclear energy which heats th central area of the Sun.

A relatively small part of particle mass i "burnt"—converted into heat or carried away b

neutrinos—in nuclear reactions. To avoid error in the difference, we should need the values of the mass of elementary particles with a high degree of accuracy (cf. Table 7).

Table 7

Masses of Particles and Nuclei (atomic units)

Particle	Symbol	Mass	Particle	Symbol	Mass
Neutrino	ν	0	Tritium nucleus	^3T	3.0170
Electron	e	0.00055	Helium-3 nucleus	^3He	3.0159
Positron	e^+	0.00055	Helium-4 nucleus	^4He	4.0028
Neutron	n	1.0090	Carbon nucleus	^{12}C	12.0005
Proton	^1p	1.0076	Oxygen nucleus	^{16}O	15.9956
Deuteron	^2d	2.0142			

Note that the mass of oxygen atom, i.e. the sum of masses of its nucleus and eight electrons, exactly equals 16. This follows from the strict definition of the mass of atomic unit accepted in physics. When mass difference is converted into energy, one should take into account that the atomic unit has a mass of 1.66×10^{-27} kg and is equivalent to the energy of 0.93 GeV or 1.49×10^{-10} J.

So let us start the first nuclear reaction. Let two protons collide. The force of electric repulsion acts between them, therefore the probability of their coalescence is very low (we shall estimate it later). Nevertheless, it is a possible event.

Thus, we have two protons on the left-hand side of the equation. What can be their probable fate? To conserve the baryon charge a deuteron, the nucleus of deuterium, the hydrogen's isotope, should be taken. To conserve the electric charge a positron should be added. The appearance of positron results in the formation of a neutrino because of the conservation of the lepton charge:

$$^1p + ^1p \rightarrow ^2d + e^+ + \nu.$$

Positron should annihilate with electron, both are excessive in plasma: $e^+ + e \rightarrow 0$. Therefore let us combine these two reactions:

$$^1p + ^1p + e \rightarrow ^2d + \nu.$$

Calculate the mass of particles having entered the reaction: $2m_p + m_e = 2.01575$. The mass of reaction products, i.e. the mass of deuteron, equals 2.0142. Thus, the mass has decreased by $\Delta m = 0.00155$ atomic units. This mass defect has been transformed into the kinetic energy of the products of reaction and the energy of electromagnetic radiation. The major portion of this energy, $E = \Delta mc^2 = 2.31 \times 10^{-13}$ J $= 1.45$ MeV, is carried away by neutrino.

Deuteron is a stable particle which does not decay spontaneously. In the Earth's water there is one atom of deuterium, heavy hydrogen, per each seven thousand atoms of the usual hydrogen. However there is a high probability for deuteron to enter a nuclear reaction with proton at the temperature of the Sun's core:

$$^2d + ^1p \rightarrow ^3He$$

Mass defect here equals 0.0059 atomic units. The energy liberated in that reaction, 8.8×10^{-13} J $= 5.5$ MeV, is all spent on the heating of the surrounding plasma. The helium isotope with mass 3 is also stable. However, its share is rather insignificant compared to the more common helium-4. On the Earth it is merely 10^{-4}%. The occurrence of this isotope is so low because at the temperatures and pressures which exist inside the stars there is a high probability for helium-3 to enter various nuclear reactions. The most probable of such reactions is the following:

$$^3He + ^3He \rightarrow ^4He + ^1p + ^1p.$$

Calculate the mass defect in this reaction: $\Delta m = 0.0138$. This means that the reaction liberates the energy equal to 2.06×10^{-12} J $=$ 12.9 MeV. At temperatures and pressures existing in the Sun's core, helium-4 nuclei practically do not enter into further nuclear reactions. The proton cycle of reactions is over.

Let us summarize the results. Six protons and two electrons were spent to produce a nucleus of helium-4, two neutrinos, and two protons. The total mass defect of the cycle equals 0.0287. This means that 0.7% of the initial mass has been transformed into energy. The formation of each helium nucleus is accompanied by the liberation of energy equal to 4.28×10^{-12} J or 26.8 MeV. Two neutrinos will carry away a part of this energy. Let us assume that they will take away approximately all the energy of the first reaction of the cycle. In that case there will be 4×10^{-12} J of energy left in each proton cycle for heating the Sun.

Other nuclear reactions are also possible in the core of the Sun. Specifically, the synthesis of helium may occur with the nuclei of carbon and nitrogen participating as catalysts. These nuclei are being transformed one into the other but finally such a carbonic cycle produces a helium-4 nucleus and the quantity of carbon and nitrogen does not change. Thus, a carbonic cycle liberates energy equal to that released in a proton cycle.

Carbon and then oxygen, and heavier elements are formed in stars as the result of the reactions $3^4He \rightarrow {}^{12}C$; ${}^4He + {}^{12}C \rightarrow {}^{16}O$ and so forth.

However the temperature inside the Sun is not sufficiently high, therefore the rate of these reactions is negligible there. But where did the heavy elements come to the Sun from? Even the quantity of lead is quite significant there. Note that the formation of elements heavier than iron ${}^{56}Fe$ is disadvantageous from the energy point of view: the mass defect changes its sign when heavy elements are formed. Recall radioactivity: spontaneous decay of the elements which are heavier than uranium. The answer is that the Sun is a second-generation star. According to modern concepts the evolution of stars passes two stages. It begins with the formation of first-generation stars out of prestar material consisting by mass of three quarters of hydrogen and a quarter of helium. These are massive stars and proton cycle reactions proceed rather fast there. Finally very little hydrogen is left in the center and combustion suspends. The star condenses, the pressure and temperature rise rapidly and helium starts to "burn". This is a critical moment in star's history. If its mass had been sufficiently

great, the synthesis of elements would be blast-like: the matter is heated to the temperatures of hundreds million degrees—the energetically disadvantageous synthesis of heavy elements takes place as well—but the star itself explodes. As this happens, both hydrogen and heavy elements are dispersed over the universe.

Note that the reactions combining helium nuclei produce elements with even charges of nuclei. Generally, the nuclear combinations of an even number of protons and an even number of neutrons are more stable at stellar temperatures. Therefore the occurrence of elements with even numbers is one order of magnitude higher than that of odd elements (consider again Fig. 24).

After the explosion of a first-generation star, the material enriched with minor impurities of practically all the elements may again assemble into stars under the action of gravitation. These very stars would be the second-generation stars which is exactly the Sun's case.

The explosion of the first-generation star occurred about 5 billion years ago to have blasted out the material from which our solar system has been formed. The majority of stars of the Galaxy are also second-generation stars. However, there are also hydrogen-helium stars in the system whose evolution has not yet approached the moment of explosion. An explosion of a star, by the way, is quite a rare event. We shall further see that such an incident requires specific conditions. Stars are stable in the case of normal burning. So let us try to answer the question:

4. Can the Sun Explode?

The stellar energy comes from nuclear reactions: the synthesis of helium from hydrogen and the synthesis of other elements from these lightest two. But hydrogen is also abundant on the Earth: the entire ocean consists of water each molecule of which contains two atoms of hydrogen. Is there a possibility that the same proton-cycle reactions are slowly developing in the Earth's ocean?

There is none. The fact is that definitely no two protons have ever coalesced into deuterium for billions of years of ocean's existence. So, why do nuclear reactions occur on the Sun and stars but not on planets?

Let us not hurry with an answer. Consider fi st how often the solar hydrogen enters into the synthesis reaction. We know the luminosity, the light power of the Sun: $L_\odot \sim 4 \times 10^{26}$ W. This power is supplied by the synthesis of helium. A formation of one helium atom releases the energy $E \sim 4 \times 10^{-12}$ J. Divide the first value by the second and you would find the number of helium atoms formed on the Sun per second:

$$\dot{N}_{\mathrm{He}} \sim \frac{L_\odot}{E} \sim 10^{38}\ \mathrm{s}^{-1}.$$

(A letter with a point above it is used to designate the rate of value alteration, its time derivative.) Certainly, this value could be calculated more accurately but for the time being we do not need a high accuracy. Thus, 4×10^{38} hydrogen nuclei, the protons, burn in a second to form 10^{38} helium nuclei and 2×10^{38} neutrinos. Neutrinos carry

away power approximately equal to $0.1L_\odot$. Therefore, nuclear reactions supply solar power equal to $1.1L_\odot$. As this happens, the Sun "loses weight" by $1.1L_\odot/c^2 \simeq 4.5 \times 10^9$ kg per second. This is quite a lot, isn't it? But the Sun itself is really huge.

Could we assume, proceeding only from this calculation, that the Sun's mass decreases? No, we could not. There is also a phenomenon resulting in the increase in the Sun's mass: the falling of comets upon the Sun. Photos were taken of several planets which were rushing into the solar disk at a speed of 618 km/s. Those were small comets and they showed their presence by glowing only few hours before they disappeared having vaporized in the Sun's atmosphere. It is difficult to estimate the mass of these comets and frequency of such falls with the number of registered cases being too little. However, the falls of comets are quite capable of compensating the radiation loss in the Sun's mass equal to 1.5×10^{17} kg per year.

Consider now how often a hydrogen atom "has to die" on the Sun, i.e. the probability for its annihilation in one second. Let us assume for estimation, that the Sun consists only of hydrogen. In that case the Sun houses $N_{\mathrm{H}} \sim m_\odot/m_{\mathrm{H}} \sim 10^{57}$ hydrogen atoms. Therefore, only 4×10^{-19} of the solar fuel burns out per second, only two protons out of each 5×10^{18} enter into a reaction per second. At such a rate even one tenth of hydrogen available on the Sun would not burn out in 10 billion years, the time comparable to the age of the universe. An energy crisis does not threaten yet our star. The low combustion rate

accounts also foı the fact that the 5-billion-yea old Sun can be still considered a young star fu of hydrogen energy.

We have calculated the probability for proto to enter into the reaction per second. Let us no try to estimate the probability for a coalescenc of two protons in one collision. For that purpos we should find out how often a proton collide with others in one second. Let us roughly assum that protons are balls with radius $r_0 \sim 3 \times 10^{-15}$ m In the central areas of the Sun where the temper ature reaches $T_s \sim 10^7$ K they move with velo cities of about $v \sim \sqrt{kT_s/m_H} \sim 3 \times 10^5$ m/ and each cubic meter houses $n \sim \rho_s/m_H \sim 10$ protons. Each proton would collide in one secon with all protons in a cylinder with a base πr and height v. The total number of protons i this cylinder is $\pi r_0^2 vn \sim 10^9$. Thus each proto collides with other protons 10^9 times per second while it enters into the reaction only once i 2.5×10^{18} seconds. Therefore the order of mag nitude of the probability for the nuclear reactio $p + p \rightarrow d$ *in one collision* is merely 10^{-28} That is for the most extreme conditions in th very center of the Sun! What can be expecte then from the Earth's ocean?

Why is the probability so low? Recall to min that protons have similar electric charge therefore, they repel. Only at the shortest dis tances of about r_0 the gain in energy from mas defect, the nuclear gravitation, becomes equa to the loss of energy in the convergence of proton The kinetic energy of a proton in the center of th Sun equals $kT_s \sim 10^{-16}$ J but protons need t spend the energy equal to $e^2/r_0 \sim 10^{-13}$ J, whic

is by three orders of magnitude more than the kinetic energy, to come to a distance of r_0 from one another. In a gas with temperature T there are always some particles with energy significantly higher than kT but the number of such particles decreases very rapidly, proportionally to $e^{-E/kT}$, with increase in energy E. Therefore, the probability for locating a particle with the energy 300 times more than kT would be about $e^{-300} \sim 10^{-130}$. This value is so small that it can be well considered as equal to zero.

If the nature of nuclear reactions were purely mechanical, the situation would be absolutely bad with *no chance for the nuclear synthesis to* occur and develop. Nature, however, has provided another opportunity. The matter is that protons do not need to climb a power "hill"with height e^2/r_0 to enter into a coalescence reaction: they can infiltrate beneath the "hill" due to tunnel effect. The passage of particles under the energy barrier is allowed in quantum mechanics.

Let two protons with kinetic energy equal to $E \sim m_H v^2 \sim kT$ collide. At long distances the potential energy of their interaction is electrostatic: $U = e^2/r$. But at the distances approaching r_0 the nuclear forces of attraction become stronger than the electric ones, therefore, at close distances potential $U(r)$ has a well the depth of which is proportional to the mass defect of the reaction (Fig. 26). According to the classical mechanics, particles cannot approach one another closer than to the distance of $r_{min} = e^2/E$ in case of such a potential. The electric "spring" is unwound and protons part again. This happens

in the majority of cases: protons elastically recoi from one another.

However at the temperature of ten millio degrees this distance approaches $r \sim e^2/kT \sim$ 10^{-12} m. It is three times more than the size o the nucleus r_0 but sufficiently small for the motio

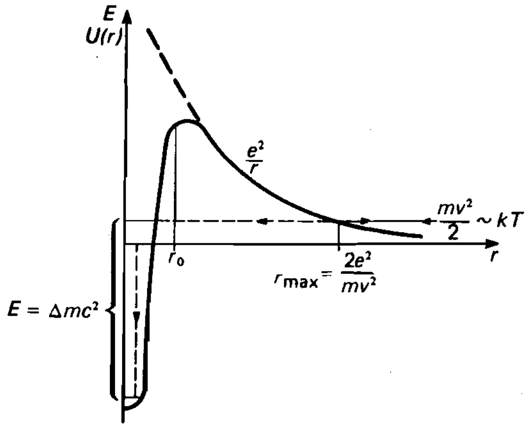


Fig. 26. Potential energy of protons approaching on another.

of particles to leave the rule of the classica mechanics and conform to the laws of the quantun mechanics.

The matter is that the motion of particle with dimensions that small cannot be describe with absolute accuracy. In case of the quantun motion one cannot precisely identify the exac location and velocity of a particle. This mean that the probability of individual events can b calculated but it is impossible to predict wha

will happen to a specific elementary particle. However in our specific case we are just interested in the probability w for the coalescence of two protons, the probability of deuterium formation, and not in the fate of an individual proton.

Proceeding from the real luminosity of the Sun we have estimated that a solar proton enters into reaction once in 2.5×10^{18} seconds. This statement sounds rather strange: one proton cannot react two times. Note, however, that the given time is longer than the age of the universe. That is why the same idea is better expressed in terms of the probability: the probability for the reaction of deuterium formation is equal to the value inverse to that time, $w \sim 4 \times 10^{-19}\ s^{-1}$. The dimensionality of w is inverse time and the sense of this value is the following: during the time t a part of protons equal to a small number wt would enter into reaction.

The probability for a nuclear reaction, w, depends upon the mean density of hydrogen and the mean energy of the motion of protons, i.e. temperature. Consider Fig. 27. The figure gives the dependence of the probability w on the temperature with the density of the material approximately equal to that in the center of the Sun: 100 g/cm^3. This dependence was calculated in conformity with the laws of quantum mechanics. This dependence actually gives the value of w approaching that calculated by the luminosity of the Sun.

If the rate of nuclear reactions in the center of Jupiter, where the temperature is about 10^5 K, were considered, it would turn out that the probability for the reaction there approaches $10^{-38}\ s^{-1}$.

This means that the rate of nuclear reactions in Jupiter's interior is by more than 18 orders of magnitude lower than that on the Sun, and the energy release on Jupiter is negligible even compared to the energy of solar light which falls on it. The more so for other planets inferior to

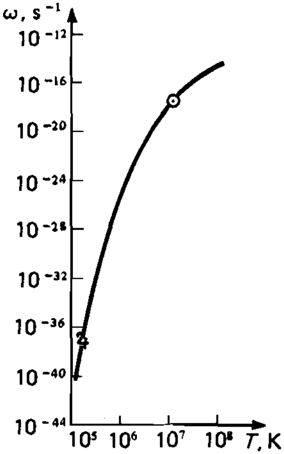


Fig. 27. Probability of nuclear reaction of deuterium formation versus temperature.

Jupiter in size. In the Earth's ocean the thermonuclear reactions of natural origin are absolutely excluded.

The knowledge of the rate of nuclear reactions brings about a new enigma. You see that the rate of a nuclear reaction rapidly increases with temperature. But the same thing is observed in some chemical reactions. For example, the reaction of water formation does not take place in a mixture of hydrogen and oxygen at room temperature, but try to put this mixture on fire and

it would explode. How does the explosion start? At the point of ignition the heating of a small volume accelerates the reaction and the energy release from it heats the mixture further. The reaction momentarily spreads over all available material—here comes the explosion.

So why doesn't the Sun explode? The procedure may seem to be practically identical there! Let us assume that the temperature has exceeded the equilibrium level at some point inside the Sun. Nuclear reactions would be intensified at that place, which would result in an even higher local superheating, bringing into a fast nuclear reaction increasingly large adjacent areas of the solar plasma. Heat removal also increases with temperature but to a lesser degree: excessive heat fails to irradiate. What happens then? The minor temperature disturbance should increase since it is unstable. Such an instability should seem to lead to an explosion, when the entire star would engage in an intensive nuclear reaction burning all nuclear fuel.

As mentioned above, stars do explode sometimes. Everything around you had once been the material of an exploded star. Flares of stars are observed even these days. It so happens that some inconspicuous star becomes suddenly, in a couple of weeks, very bright, the luminosity increases by millions of times. Astronomers call such a star a nova and the observed event is termed a nova flare. After the flare the star slowly—in several months—returns to the more or less initial state.

The even grander but more rare events were termed supernova flares. The last flares of super-

novas were observed in our Galaxy in 105 1572, and 1604 (those were certainly the year when the light from explosions reached us). Th last two flares had revived the interest of hu manity to astronomy, soon after the telescop was invented. But as ill luck would have i there have been no supernova flares in our Galax since then.

After the explosion of a supernova, the luminos ity also decreases rapidly, but the star looses it old appearance. A fast-rotating neutron sta the pulsar, takes the place of the flared super nova and all other material is scattered with hig velocity. The size of neutron stars is merel about 10 km but their mass approaches that o the Sun. Their gravitational field is so powerfu that under the action of immense pressures th electrons of all atoms are pressed into nuclei the protons of nuclei turn into neutrons.

Yet there is no threat of such terrific transform ations for our star. The thermonuclear reaction progress steadily. Very special conditions ar required for a star to explode. That is what pro vides the stability of stars. If for some reaso extra energy is liberated inside a star, the pressur at that place would be equalized and the sta itself would slightly expand much earlier tha intensive nuclear transformations would hav a chance to start. As this takes place, the sta radius increases. The temperature of the sta interior, as the formula indicates, is inversel proportional to the radius of the star, therefo it should cool down. Thus, the liberation o energy, the heating of the star, results in a de crease in its temperature. Stars seem to determin

their size themselves. The composition of stars being similar, the luminosity and radius depend exclusively on mass.

The main reason for the star stability is the slow rate of nuclear burning and heat removal compared to the time of pressure equalization. Pressure disturbed by some effect tends to return to equilibrium value at the rate of the sound velocity. The sound velocity approximately equals the velocity of heat motion of atoms $u \sim \sqrt{kT/m_{\mathrm{H}}}$. Therefore, even the greatest in size pressure disturbances inside the Sun can be equalized during the time

$$\tau_1 \sim R_{\odot}/u \sim 25 \text{ min.}$$

The specific time of energy heat removal from the central areas of the Sun is much longer. Let us estimate the time during which the energy, released by nuclear reactions, reaches the surface of the Sun. Assume that reactions in the center have ceased. After what time could that be detected by the cooling of the surface? The specific heat of the solar plasma equals, by the order of magnitude, k/m_{H}. Therefore, the total store of the Sun's thermal energy is $E_{\mathrm{t}} \sim m_{\odot} kT_s/m_{\mathrm{H}} \sim 3 \times 10^{41}$ J. Divide this energy by luminosity, the thermal power of the Sun, to estimate the time of its thermal cooling:

$$\tau_2 \sim \frac{E_{\mathrm{t}}}{L_{\odot}} \sim 3 \times 10^{14} \text{ s} \sim 10^7 \text{ years.}$$

It turns out that the Sun would shine for a full geological epoch even without the nuclear energy. However, we know that the Sun is much

older. The time of burnup of solar hydrogen also significantly longer, not less than 10^{10} yea

The absolute value of the gravitational ener of the Sun turns out to be about the same or of magnitude as its thermal energy E_t. Gra tational forces are forces of attraction, theref the gravitational energy is negative. For t bodies with masses m_1 and m_2 removed to t distance R from one another it equals $-Gm_1m_2/$ To calculate accurately the intrinsic gravitatio energy of the Sun, one should take into accou the distribution of density in it, i.e. the depen ence of density upon the distance to the cent But if we consider only the order of magnitu of the energy of gravitational attraction, could write the formula without a numeric coefficient:

$$E_{gr} \sim -\frac{Gm_{\odot}^2}{R_{\odot}}.$$

The modulus of this value turns out to be equ to 4×10^{41} J; it is actually of the same order magnitude as the thermal energy E_t.

From the law of conservation of energy it fo lows, that a decrease in the size of a star resul in the liberation of thermal energy. W. Kelv and H. Helmholtz, the physicists of the la century, believed, that it was exactly the gra tational energy which was transformed into radi tion by compression of stars. They were not awa of thermonuclear energy, therefore their calc lations showed, that the Sun was merely 10 milli years old.

However, the role of the gravitational ener in the evolution of stars is also very importa

The decrease in it due to compression heats the material in the process of formation of stars up to temperatures, at which the nuclear burning of hydrogen starts. From that point on the stability of stars is supported by the balance of the thermal and gravitational energy.

5. The Sun's Atmosphere

When is a physical problem complicated? Maybe, when its object differs significantly from the conventional scale? The answer is "no", since the structure of the solar interior is rather simple. The physics of atomic scale is also not difficult to comprehend. A problem becomes complicated when it involves several competing physical phenomena developing simultaneously, when none of the parameters can be neglected.

That kind of situation had not existed in physics for all times. It was time, when the principal laws of nature were unknown and complete classes of phenomena were complicated and inexplicable. Presently we know the fundamental laws of physics: the theory of gravitation, the theory of electromagnetic phenomena, the quantum mechanics and statistics (the theory of systems composed of a large number of particles). There are significant gaps only in high-energy-particle physics. But we have an impression that this area will be also clarified soon.

Is it, however, possible that other fundamental physical principles will be discovered? Yes it is, but not in the well-investigated areas of science. Even the potential unification of all interactions into one system will not cancel the accumulated

knowledge. Perhaps, only biophysics has opportunity to discover the laws of nature, oth than the above mentioned. Yet the number simple, strictly identified physical facts is n so far sufficient to establish fundamental la of biology, which are not the consequences the known laws of physics. At the same time the are numerous phenomena around us which sti await their explanation. Why is it so?

Here is a characteristic example. Let us t to analyze the near-surface structure of the Su The Sun's surface can be observed by eye an examined with the help of the instrument These days nobody doubts, that the fundament laws of physics hold true for the solar surfac they also work well in the depths of the Su Nevertheless we have so far failed to explai authentically many of the effects observed on th Sun.

Imagine that you can see the Sun's surface i monochromatic light, viz. light with a singl wavelength. Then it turns out that at some wave lengths the Sun looks completely different tha in white light (Fig. 28). What is more, a chang in wavelength of only a hundredth of a nan meter, for example, from 656.47 nm in any direc tion, radically changes the observed pictur Why is this so?

We have in fact not chosen this wavelength random. The light with that wavelength has specific property: it can cause a quantum trans tion of a hydrogen atom from one state to anothe more energetic and excited. And conversely, hydrogen atom can transit from the excited sta to a less energetic one having emitted a ligh

quantum, a photon with exactly that wavelength. This wavelength can be expressed in terms of the fundamental constants:

$$\lambda_{2,3}=\frac{4\pi c\hbar^3}{m_e e^4}\cdot\frac{1}{\left(\frac{1}{2^2}-\frac{1}{3^2}\right)}=6.565\times10^{-7}\text{ m}.$$

The formula is written in that rather unusual manner in order to distinguish numbers 2 and 3;

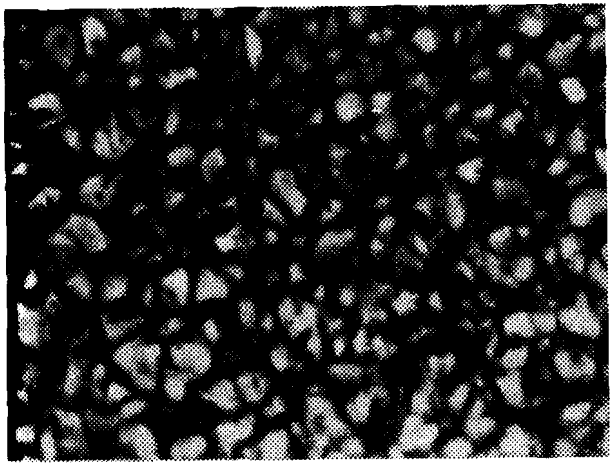


Fig. 28. Photograph of the solar surface in a wide range of visible light.

the light of that wavelength is irradiated or absorbed in transitions of a hydrogen atom from its second level to the third or back to the second. Any two integral numbers could substitute 2 and

3 in the formula to produce a multitude of wa lengths of light which interacts with hydrog atoms in a special manner. However, the maj portion of this radiation does not get into t visible band of the spectrum.

The reason for the Sun to look in the hydrog light with 656.5 nm wavelength quite differe from its appearance in the light of other no singular wavelengths is that different lig filters make a camera to register the distributio of luminance at various altitudes of the sol atmosphere; it is like making slides of the sol atmosphere at various levels. Hydrogen is t most common element on the Sun. Therefor radiation which actively interacts with hydrog entangles in the acts of absorption and irradi tion at high altitude, where the gas is sufficient rarefied and transparent for the light of oth wavelengths. The deeper layers of the solar a mosphere are visible in the usual light.

We have asked why the Sun looks differen if viewed in different light and the answer seem to be quite adequate. But we have not stoppe for a moment to think of the reason for the Sun atmosphere being so different at various altitude to think of what happens with a rise by th thousand kilometers (a mere one seven-hundredt of the solar radius) which separates the altitud given in Fig. 28. Is it not an example demonstra ing that knowledge, the same as the Sun, ca have various depths of cognition?

Figure 28 gives the lower layers of the Sun atmosphere, its main horizon of radiation, th photosphere. It looks like a layer of spille grain. These formations are called granule

Granules are the observable demonstrations of convection in the photosphere. The upward motion of the heated gas occurs, as a rule, in the bright centers of granules while at the edges of granules the gas, cooled by the energy losses for radiation, descends. The observable contrast in luminance results from that very temperature difference. The dimensions of granules vary between 200 and 1300 km; the solar disk houses about a million of them. The granules cannot be distinguished by a naked eye. Just as in the case of a usual convection, the convection layout is not stationary on the Sun: an individual granule lives for about 10 minutes after which its boundary becomes indistinct and the place is taken by new granules.

The extended structures have cross dimensions approaching those of granules. They are called fibrilles when they are observed near the center of the solar disk. The same formations observed on the limb (the edge of the solar disk) are called spicules, the word stemming from Latin *spiculum* which means "edge", "sharp end", "sting". The area of altitudes, where fibrilles and spicules are observed, is called chromosphere; it is located at 600 to 2000 km above the level shown in Fig. 28.

What physical features does the observable surface of the Sun have? Firstly, the surface is visible, we see the boundary of the solar disk. This means that the boundary sets a limit beyond which the solar energy, the radiation, can leave the star for good. The radiation's departure results in a significant cooling of the Sun's sur-

face, the photosphere, compared to the deepe
layers.

The temperature of the photosphere is sufficient-
ly low (although not lower than 4000 K) for ion
of hydrogen, helium, and other elements to cap-
ture electrons and turn into neutral atoms. Thi
process is not completed, there remain stil
enough free electrons and ions to support a hig
electric conduction in the environment. At th
same time even simple molecules are bein
formed in small quantities: CO, H_2, CH, CN
Gases of the solar atmosphere absorb the radiatio
of the lower, hotter layers, their atoms an
molecules are excited, while absorbing photons
and then transit to an unexcited state by emittin
photons. The density of atoms in the photospher
rapidly decreases with height, diminishing twic
per each 130 km. For each frequency (or wave
length of light) there is an altitude, at which th
gas turns transparent and the radiation of thi
frequency leaves the Sun.

We have seen from the example of hydrogen
that the closer the frequency of light to the fre-
quency of atom's quantum transition, the stron-
ger the light of this frequency interacts with th
material and, therefore, the higher position i
taken by the layer of the solar atmosphere, wher
the radiation of this frequency is formed. Ye
in the photosphere the temperature falls wit
height and the intensity of radiation falls rapidl
as the temperature goes down. That is why th
radiation of the Sun is usually weaker at th
frequencies, which correspond to quantum tran-
sitions of atoms and molecules. These dark band
in the solar spectrum are the Fraunhofer lines

which were mentioned in the beginning of the chapter.

Consider now again Fig. 23, especially the most detailed part of the spectrum, Fig. 23*c*. The atom of each of several dozens of elements, which are present on the Sun in significant quantities, has about a hundred of quantum levels of energy excited at solar temperatures. Not every two levels can emit a photon when a transition between them occurs—there are several quantum rules of forbidden transitions—but nevertheless the total number of possible transitions of the solar atoms and molecules, which affect the resultant spectrum, reaches several millions! The result is about 20 thousand maximums and minimums observed in the solar spectrum.

As this takes place, the spectrum of the Sun changes throughout its surface, which can be seen even in photographs. In addition, irregular time changes of the spectrum occur at every point of the solar surface.

All those changes (in space, time, and light frequency) carry an immense quantity of information. Imagine that a problem was set just to record, to register in the memory of computers the entire variety of solar spectra reaching the Earth, all that can be in principle observed with the help of modern instruments and record techniques without any explanation or investigation into the results of observations. It turns out that the volume of this information exceeds by 4 to 5 orders of magnitude, by tens and hundreds thousand times, that of the total information which we get from the rest of the universe!

Information is a mathematical notion. If the nature of physical processes is clear, the relations between phenomena are explained, there is no need to record all things, which can be observed. Only the unpredictable and incomprehensible facts should be memorized. We are presently well advanced in the physics of the phenomena in the solar atmosphere, but there are still many things, which remain so far inexplicable.

The analysis of radiation spectra throughout the entire band of electromagnetic waves makes it possible to construct the models of the solar atmosphere for areas of the surface with different luminance and structure. The chemical composition of the atmosphere can be determined by the wavelengths of the Fraunhofer lines and their relative intensities. The forms of lines permit to find the temperature, density, and the degree of ionization of the solar atmospheric gas at various altitudes. The minor shifts of lines due to the Doppler effect permit to calculate back the velocities of gas motion. From the splitting of lines and polarization of the radiation the value of the magnetic field can be calculated for various areas of the solar surface.

Cycles of continuous observations for periods of several days (a station located near the South Pole of the Earth was used for that purpose) allow to find the spectra of the waves running over the Sun's surface. As theory has predicted, these waves turned out to be a complicated combination of sound waves and surface waves. The surface waves of the Sun are, to some extent, like the waves of terrestrial seas but their real nature is much more sophisticated: the water

surface of the Earth has a distinct boundary with light air, while on the Sun the density of gases decreases continuously with height.

The Sun rotates. This rotation occurs in the forward direction, i.e. in the direction coinciding with the revolution of the planets. However, the axis of the Sun's rotation is not strictly perpendicular to the ecliptic but makes an angle of 7°15′ with the normal to it. The North Pole of the Sun can be observed on the Earth between June 7 and December 7, during the other half of the year the South Pole is observable.

We assess the rotation of the Sun by the regular displacement of the points of its surface. However this gaseous sphere rotates not in the manner of a homogeneous solid body, not like the Earth: a point on the equator of the Sun's surface makes a full turn in 25 days, while the period of rotation near the poles is 35 days. Such a heterogeneous rotation is called differential in contrast to solid-body rotation in which the angular velocity of all points is the same. The angular velocity of the Sun's rotation also changes with depth, but the actual mechanism of this change is yet unknown. One thing is clear: the convection of the Sun, its magnetic fields, and differential rotation are related and interact with one another.

The convection of an ionized gas in a magnetic field is much more complicated than the usual one, since the electric currents, generated in this case, produce themselves magnetic fields. Apparently, something unusual and incomprehensible takes place somewhere at the boundary between the photosphere and the chromosphere as a result of a complicated interaction of the

quantum radiation, waves, coming from convective depths, and magnetic and electric fields. This is evident because the temperature of the solar gas, which falls on the way from the center of the Sun to the photosphere (a natural thing when the distance to the source of energy increases), suddenly starts to rise (Fig. 29).

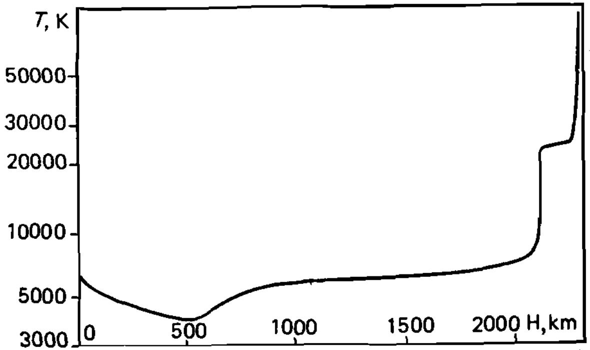


Fig. 29

After the temperature reaches its minimum at an altitude of about 500 km, where it falls down to 4400 K, it rises first slowly (in the chromosphere) and then, in the so-called sheath between the chromosphere and the Sun's corona, the temperature rises by two rapid jumps increasing twice with an altitude increase of mere ten kilometers! From that level the Sun's corona stretches.

The corona, the temperature of which reaches 1.5 to 2 million kelvins, makes a halo of irregular and inconstant form easily observable during total solar eclipses. Individual rays reach out to

the distance of 1.5 to 2 solar radii. "Besides the peculiarities of this atmospheric "average" (the so-called "quiet Sun"), various surprising phenomena take place in the solar atmosphere and await still their explicit explanation. In addition to the granules, supergranules, and spicules characteristic for the normal (undisturbed) solar atmosphere, certain "disturbances", i.e. inhomogeneities with relatively short lives, are also observed there. The major of those include sunspots, faculae, chromospheric bursts, prominence, coronal rays and holes." That was a quotation from a scientific paper on the hydrodynamics of the Sun.

One cannot help recalling "Solaris" by Stanislaw Lem, where the behaviour of a live ocean of a fantastic planet was described: "No terms can give the idea of what is happening on Solaris. Its *treemountains*, *longlets*, *mushroomies*, *mymoids*, *symmetriads* and *asymmetriads*, *spiners* and *fasters* sound terribly artificial but they give an impression of Solaris even to those who have seen nothing but vague photos and extremely inadequate films." The analogy is also traceable in the name given by Lem to the planet.

The corona gives birth to the solar wind. The escape velocity $v_2 = \sqrt{2Gm_\odot/r}$ decreases with the distance from the Sun. It equals three hundred kilometers per second already at the distance of $r = 4R_\odot$. The average particle velocity at the corona temperature of about 1.5 million degrees approaches 100 km/s and many protons move thrice faster. Therefore the velocities of many plasma particles are sufficient for escape. At the same time the plasma density rapidly decreases

with the distance from the Sun. Beginning with the distance between 2 and 3 solar radii the collisions of particles become so rare that these particles can easily leave the Sun and rush into space. This is the solar wind.

A clear demonstration of the latter, the tail produced when that wind blows over comets, is known from ancient times. However there was only a hypothesis on the existence of a flux of charged particles emitted by the Sun before the solar wind was actually discovered by spacecraft. At the distance of the Earth's orbit $a_\oplus$ the mean density of the solar wind is $n = 5$ particles per cubic centimeter and its average velocity is $v = 320$ km/s.

Let us estimate the mass lost by the Sun per second due to the solar wind. This will be the mass density of $0.5\ nm_H$ multiplied by the velocity v and the surface of the sphere $4\pi a_\oplus^2$ with the radius of the Earth's orbit:

$$-\dot{m}_\odot \simeq 2\pi a_\oplus^2 n m_H v \simeq 4 \times 10^8 \text{ kg/s}.$$

It is known that this value is only by one order of magnitude lower than the Sun's loss of mass due to mass defect in the course of nuclear reactions: 4.5×10^9 kg/s.

Both the intensity of the solar wind and the corona size alter significantly according to the solar activity. Another apparent demonstration of the changes in the upper layers of the Sun are the sunspots. The more spots appear on the Sun and the larger they are, the farther the solar corona stretches and the stronger the solar wind.

Large spots on the Sun are observable even to the naked eye. They were mentioned already in an

cient Chinese chronicles. Yet the history of the really scientific investigation of sunspots started with the invention of the telescope. In 1611 sunspots were described simultaneously by several European scientists including Galileo.

Since then a number of important regularities of the solar activity were established. It was noted by the middle of the last century that the number of sunspots alters almost periodically and the period of this alteration approximates 11 years. These changes in the Sun's activity are observed until the present. It is known that sunspots are always accompanied by faculae around them. Magnetic fields in sunspots have been measured: their value may exceed that of magnetic field outside the spots by a thousand times. It is known that the polarity structure of magnetic fields in spots alters with the period of 22 years, which is twice longer than the period of the number of spots. Many other regularities have been discovered.

Unfortunately, the major part of our knowledge of the solar activity is descriptive. The relationships between some of the observed phenomena can be explained but we have failed so far to develop a comprehensive theory of the solar activity. There is not even a generally accepted theory on the qualitative level with numerical estimates. Such a primary theory should provide arguments for a well-grounded estimate of the period of the solar activity equal to about ten years. Alas, this has not been done yet.

In such a situation people would like to apply observations made in the last two centuries to longer periods. The erroneousness of that approach

11-01137

to the solar activity can be demonstrated by the following example: between 1645 and 1715 there were almost no spots on the Sun! Before 1645 the solar cycle did not show the 11-year regularity. When E. Maunder, the superintendent of Greenwich observatory, brought these facts to light from old publications in 1890, nobody believed them. It was much easier to explain the absence of sunspots by a chance or irregularity of observations than to assume a real interruption of the solar cycle. Yet the existence of this 70-year break in the solar activity (called Maunder minimum) can be hardly doubted.

It is confirmed both by the reaction of European scientists who considered indvidual sunspots as extraordinary events at that time and the small number of auroras registered during that period. Auroras are known to stem from solar wind intensity splashes. Finally, annual rings of trees corresponding to the time between 1650 and 1720 show a 10% increase in the contents of radioactive carbon ^{14}C. A small quantity of this isotope is formed from nitrogen in the upper layers of the Earth's atmosphere under the action of cosmic rays. Then ^{14}C is assimilated by vegetation in the course of photosynthesis. The radioactive carbon ^{14}C with half-life of 5570 years decays to form the conventional nitrogen ^{14}N. The concentration of ^{14}C allows archeologists to date wooden articles rather accurately.

Cosmic rays are a flux of high energy particles falling to the Earth from all sides. Nuclear reactions in the high layers of atmosphere are caused by cosmic rays not of the solar but of the galactic origin. Nevertheless the intensity of these cosmic

rays is related to the solar activity: the more spots on the Sun, the weaker the flux of cosmic rays. The relationship is realized through the solar wind.

The solar wind compresses the force lines of galactic magnetic fields at the distance of about 10^{13} m, which is approximately fifty radii of the Earth's orbit. Cosmic rays propagate mainly along the magnetic field. Only the most energetic particles of galactic cosmic rays can penetrate this magnetic bubble around the Sun. The solar wind covers the distance to its boundary in six months or a year. The Sun being quiet, the solar wind is weaker and the boundary of the solar magnetosphere moves closer and becomes less dense. The result is an increase in the intensity of cosmic rays reaching the Earth and the contents of ^{14}C in atmospheric carbon dioxide.

The period of quiet Sun in the usual 11-yerar cycle is not sufficient for the accumulation of the radioactive carbon in a quantity ample for tracing the difference from the years of active Sun by annual rings of trees. But this difference became quite detectable during the 70-year-long Maunder minimum.

The Sun's thermal radiation is constant. The solar activity alters only the short-wave, nonthermal part of radiation at wavelengths less than 100 nm. The second, lesser, maximum can be seen in Fig. 23. However, less than 1% of the Sun's total luminosity falls on this area. As we shall see further, the short-wave part of the solar radiation does not pass through the upper layers of the Earth's atmosphere. Therefore the solar activity practically does not change the heat

flux coming to our planet and almost does no affect the Earth's weather. The alternate short wave radiation changes significantly only th state of the external shell of the Earth's atmo sphere.

Unfortunately, the physical origins of the so lar activity and other phenomena in the Sun' atmosphere are still far from being absolutel clear but the Sun's effect on our world is im mense. We shall further return to its minor part the thermal effect of the Sun upon the Earth. For the time being just listen to poet's word:

Almighty Sun, the nature's vital spark and grace,
Eternal light and holy face of God,
Thou blow life into air and earth, thou make wate blaze,
Thou is time and matter's Lord!

Alexander Sumarokov, 1760

Chapter 4

Atmosphere and Ocean

1. The Composition of the Atmosphere

Have a look at the blue sky: this is atmosphere. Take a deep breath: this is also atmosphere. But what is atmosphere from the point of view of physics? What accounts for its composition, pressure, and temperature at different altitudes? What is the cause of its motion, the cause of winds? How do the winds produce ocean currents? Let us try to supply brief answers to these questions.

You know that 78% of the Earth's atmosphere is taken by nitrogen, 21% falls on oxygen, and argon accounts for 1%. There are also minor admixtures of carbon dioxide, water vapours, and quite negligible quantities of neon, helium, krypton, and hydrogen. But let us try to comprehend why the upper shell of this planet consists of these very gases and world ocean water. The atmospheres of other planets of the solar system are totally different and they have no ocean in the conventional sense of the word. Carbon dioxide prevails on Venus and Mars, the major atmospheric components on giant planets are helium, hydrogen, methane, and ammonia. The Moon and Mercury do not have any atmosphere at all. Why is it so?

The atmospheric composition is determined first of all by the geological history of the given planet. The planets of terrestrial group, which

were formed by way of amalgamation of mineral solid particles, apparently did not have any atmospheres at the initial stage. The primary planetary material was gradually compressed by gravity, the planets took a spherical form while their interior was heated. Chemical reactions were induced in the primary material under the action of high temperatures and pressure. The heavier reaction products were sinking down, thereby forming the nucleus, the lighter ones formed the outer solid crust of planets and the gaseous products of reactions made up the first atmosphere. The Earth's atmosphere had accumulated so much water vapour that the major part of it could not fail to condense. Thus, the primary ocean was formed.

Do not think that all those things had happened once and for good in ancient times. This process is still taking place on the Earth although not so intensively as at the beginning of the evolution. The Earth's crust is still being regenerated and the volcanos are still releasing a lot of water vapour and carbon dioxide into the atmosphere. In addition to them there are also sulphur dioxide, hydrogen chloride and other unpleasant gases. Why are they absent in the normal atmosphere?

The answer is almost evident. All gases of the atmosphere should be in chemical equilibrium both with one another and the ocean, as well as with the material of terrestrial rocks. Therefore acid oxides released by volcanos dissolve rapidly in the ocean and form acids. These acids by interacting with the major oxides of the Earth's crust produce salts. The soluble salts are dissolved

in the ocean, the insoluble salts build up sedimentary rocks.

An attentive reader could have certainly noticed a weak point of the above arguments. Oxygen! This gas is absent in volcanic gases and atmospheres of other planets.

The major source of the Earth's oxygen is vegetation. The chlorophyll of plants processes carbon dioxide under the action of solar rays. The carbon from carbon dioxide enters into organic compounds of plants and oxygen is released into the atmosphere. There is however another source of oxygen on this planet. To understand how it acts, one should first clear up the question: what holds the atmospheres of planets, why don't their gases volatilize into space?

The air pressure at the Earth's surface is known to be equal to $p_0 = 1.013 \times 10^5$ Pa. This means that a total force of $4\pi R_\oplus^2 p_0$ acts upon the entire Earth's surface with the area $4\pi R_\oplus^2$. The primary source of this force is certainly gravitation. According to Newton's second law this force equals the mass of the Earth's atmosphere multiplied by the free fall acceleration g. Hence it is not difficult to calculate the mass of the Earth's atmosphere:

$$m_A = \frac{4\pi R_\oplus^2 p_0}{g} = 5.3 \times 10^{18} \text{ kg}.$$

This accounts, as you can see, for almost a millionth part of the entire Earth's mass. It is also interesting to compare the mass of the atmosphere to the mass of water on this planet: 98% of water is taken by oceans, about 2% is contained in glaciers, mainly in Antarctica and Greenland,

while the mass of freshwater bodies and water vapours is relatively insignificant. The total mass of the Earth's water equals 1.4×10^{21} kg, which exceeds 266 times the atmosphere's mass.

Free fall acceleration does not merely produce atmospheric pressure near the Earth's surface. It also impedes dissipation, the scattering of atmospheric gases into space. Let us compare the velocities of molecules of different gases at the temperature, say, of 300 K with the Earth's escape velocity $v_2 = \sqrt{2gR_\oplus} = 11.2$ km/s. It is clear that if the thermal velocity of gas molecules $v \sim \sqrt{kT/m}$ approaches the escape velocity, this gas would not stay in the atmosphere. In fact, the thermal velocity of hydrogen molecules equals 1.1 km/s, that of helium molecules is 0.8 km/s, and the average velocity of nitrogen and oxygen molecules approaches 0.3 km/s.

At first sight there is nothing wrong. The velocities of gases turned out to be less than the escape velocity. This means that the Earth can hold any gas in its atmosphere. But in fact the atmospheric gases do volatilize into space but the process is very slow. This happens because *in the top layers of the atmosphere the temperature is much higher than near the surface.* We shall see further that it reaches 1200 K. This *means that the velocities of gas molecules are twice higher there than the above estimates.* Besides, we have estimated only the average velocities of the molecules. These are the velocities of the absolute majority of molecules in the state of thermal equilibrium. But a minor part of them move at a speed by far exceeding the average

value. These fast molecules can leave the Earth for good.

You have seen that at the same temperature the highest velocities are demonstrated by hydrogen and helium molecules, i.e. gases with the minimum molar mass. Consider Fig. 30. It gives the masses of various gases of the Earth's air

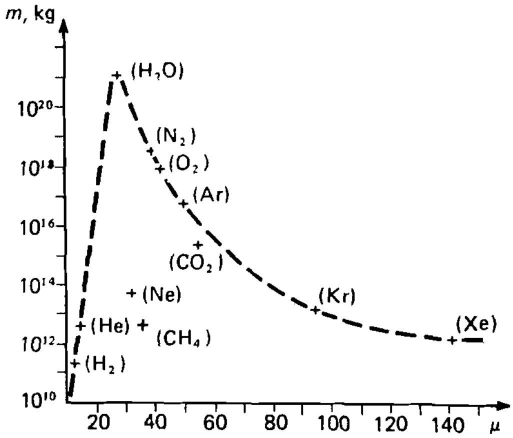


Fig. 30. Masses of gases of the terrestrial atmosphere and ocean water versus their molecular mass.

and the total mass of water on the Earth as a function of their molecular mass. You may see that the majority of points are located close to the curve which reaches its maximum at the molecular mass of water. The relative content of various elements on the Earth depends first of all on the quantity in which they are produced by nuclear reactions on stars. Therefore Fig. 30 should be

compared to Fig. 24 which gives the occurrenc of elements in space. You see that there is ver little helium in the Earth's atmosphere while o the Sun and giant planets it accounts for on quarter of their mass. The reason for this is th the Earth's gravitation is not sufficient for a sa fe containment of helium and hydrogen.

The density of these gases is also lower tha the others of their kind: they always tend to ris upwards in the atmosphere. The Earth's at mosphere turns out to be homogeneous, i.e. wel mixed, only below 90 km. This part of it is cal led homosphere. The place above it is taken b heterosphere, the part of the Earth's atmosphe with variable composition. The concentratio of helium and hydrogen in heterosphere increase with altitude. Above 700 km the Earth's atmos phere consists practically only of these gase It is mainly helium and hydrogen which volati lize into space.

However, the concentration of helium an hydrogen in the homosphere is extremely low. Bu where do they come from to the heterospher Presumably, their quantity should continuosl decrease because of volatilization.

The matter is that there are processes suppor ing the content of these light gases in the at mosphere. Helium is formed in the Earth's cru in decay reactions of heavy radioactive element As regards hydrogen of the upper atmosphe it is formed from water! At the altitude of 3 to 50 km H_2O molecules disintegrate into hydr gen and oxygen under the action of the ultra-vi let solar radiation. Therefore the volatilizati of hydrogen into the outer space results in a d

crease in water on the Earth and an increase in oxygen content in the atmosphere.

Each second approximately 1 kg of hydrogen escapes from the Earth's atmosphere into the outer space. Is it much or little? Let us estimate how long would oceans and glaciers last at such a rate. 1 kg of hydrogen is contained in 9 kg of water. Divide the ocean's mass by this rate of decrease and the result will be that the store of the Earth's water would suffice for 1.5×10^{20} s which is 5000 billion years. Thus, the ocean can be considered inexhaustible since the Earth's age is "merely" 4.5 billion years.

Estimate now the amount of oxygen formed during the time of the Earth's existence. This will be a very rough estimate because we cannot really assume that the Sun had been shining just as now during all that time. Yet let us try: 8 kg per second in 4 billion years results in 10^{18} kg of oxygen which is exactly the quantity presently contained in the atmosphere of this planet, one fifth of the mass of the atmosphere.

This coincidence should not be overestimated. It has been taken much more oxygen than its present content in the atmosphere to bring the Earth's chemical equilibrium to its present state, to oxidize methane and ammonia of the primary atmosphere, to oxidize all rocks of the Earth's crust. This would have been impossible without vegetation. The plants produce about 10^{14} kg of oxygen per year, i.e. 3×10^{6} kg per second. This is much more than the dissipation of hydrogen into outer space produces. However, presently the content of oxygen in the atmosphere does not increase: all oxygen generated by vege-

tation is consumed by living beings and spen on the oxidation of volcanic gases, combustio and decomposition of dead plants.

The dissipation of gases into space qualitative ly determines the composition and masses o the atmospheres on other planets of the solar sys tem. However, to comprehend the origins of th variety of their atmospheres one should first le arn to determine the temperatures of planet at mospheres.

2. Thermal Equilibrium of Planets. Their Atmospheres and Oceans

The surfaces of planets and their atmospheres ar heated by the solar radiation. The Sun heated b thermonuclear reactions irradiates uniformly i all directions the light power $L_\odot = 3.83 \times 10^{26}$ W. How much of it falls on the Earth? It is easy to see that this part equals the ratio o the area of the Earth's disk, $\pi R_\oplus^2$, to the surfac $4\pi a_\oplus^2$ of a sphere with the radius of the Earth' orbit, $a_\oplus$. The Earth gets merely a half of on billionth part of the entire solar radiation. Thi part multiplied by the Sun's luminosity $L_\odot$ gives the light power received by this planet

$$P = L_\odot \frac{R_\oplus^2}{4a_\oplus^2} = 1.75 \times 10^{17} \text{ W}.$$

What are the consumers of this power? A par of it is reflected by the Earth itself back into spa ce. Other planets and the Moon are visible on th night sky just due to reflected light. In the sam way one could see light reflected by the Eart when moving away from it into outer space. Th spectrum of this light differs from that of th

source, i.e. the Sun, because the Earth's atmosphere and ocean absorb red rays better than blue rays. That is why our planet looks blue from space.

The share of reflected luminous energy is called albedo and is designated by A. This Latin word stems from *albus* meaning "white". Albedo is a kind of a degree of whiteness. The accuracy of our knowledge about the albedo of our own planet was rather low up to a recent time. The Earth was believed to reflect into space between 30 and 40% of light falling on it. Recent measurements made from satellites indicated the value of albedo equal to 28%. But where does the remaining part of power, $P(1 - A)$ go? It is clear that this very power is the reason for the climate on the Earth being warm compared to surrounding space. However, the solar radiation comes to the Earth continuously and if there were no heat removal, temperature on the Earth would rise continuously and very fast. Therefore, heat removal does exist. It is easy to guess that it is realized through the same physical process as the radiation of the Sun itself: the thermal radiation. Amazingly, the Earth and other planets are also sources of radiation. But the temperature of this radiation is low, so the maximum of its spectrum falls on wavelengths far from the visible range. Thermal radiation of the Earth and planets occurs in infrared range, it is invisible to eye.

Let us calculate the temperature of the equilibrium thermal radiation of this planet taking its albedo equal to 28%. We have to equalize the thermal flux of the radiation absorbed by the Earth $P(1 - A)$ to the Earth's total thermal radiation flux. On the average, the planet

irradiates like a black body with radiation temperature $T_\oplus$ which we are going to find out. According to Boltzmann principle, the density of radiation flux is $\sigma_B T_\oplus^4$ and, to determine the total flux, one should multiply this value by the area of the Earth's surface $4\pi R_\oplus^2$. Therefore

$$4\pi R_\oplus^2 \sigma_B T_\oplus^4 = (1-A)\, L_\odot \frac{R_\oplus^2}{4a_\oplus^2}$$

$$= (1-A)\, 4\pi R_\odot^2 \sigma_B T_\odot^4 \frac{R_\oplus^2}{4a_\oplus^2}.$$

Extracting hence the fourth root we get:

$$T_\oplus = T_\odot \sqrt{\frac{R_\odot}{2a_\oplus}}\,(1-A)^{1/4} = 257\ \text{K}.$$

At that temperature the maximum of the thermal radiation falls on the wavelength $\lambda = 1130$ nm which is 22 times longer than the wavelength of the maximum of the solar spectrum.

Thus the Earth's temperature amounts to 257 K or -16 °C. Quite a frost, isn't it? But if we assumed the albedo to be equal to 0.4, the resulting temperature would be still lower -28 °C! What is the matter, however?

We know that the mean annual temperature in the Earth's middle latitudes is above zero and a large portion of the Earth's surface is taken by tropics where temperature rarely falls below 25 °C both in summer and in winter. Is it possible that the Earth has a source of heat other than the solar radiation? For example, it could be heated by a radioactive decay of heavy elements in its interior or by heat stored deep inside the planet at the moment of its formation.

Such a source of heat does actually exist. However, estimates given in the section "Why is the interior of planets hot?" demonstrate that the heat flux from the depths is by 3 orders of magnitude weaker than the heat which the Earth gets from the Sun. The heat flux from the Earth's depths taken alone could only maintain on the Earth's surface a temperature by an order of magnitude lower than the present one, i.e. merely 35 kelvins. This is why the climate of our planet is practically not affected by the heat of the Earth's depths which holds true even for the polar night, when there is no direct supply of the solar power, or for areas near the boundaries of drifting plates, where the heat flux from the depths is relatively powerful.

Thus the discrepancy between the Earth's radiation and surface temperatures cannot be accounted for by the heat flux from its interior. We have to look for another reason. Here it is. The fact is that heat is irradiated into space, as a rule, not by the Earth's surface itself but by the Earth's atmosphere, the air shell around the planet! At first sight it may seem strange since the air is transparent and should not impede the radiation of the Earth's solid surface and ocean. But the spectrum of radiation at the temperature of the Earth's surface of about 300 K lies within the far infrared zone. Therefore we cannot judge the radiation-transmitting properties of materials only by our own sensations. Measurements made with the help of infrared spectrometers indicate the following. The major air components—nitrogen, oxygen, and inert gases—are also transparent in the infrared band. A

ray coming through the entire thickness of tl atmosphere composed only of these gases wou weaken insignificantly even if this ray is an infr red one. However, the carbon dioxide CO_2 an water vapours H_2O, which are present in the a mosphere in minor quantities, absorb the infr red radiation so intensively that they become d cisive factors of the transparency of the Earth atmosphere in infrared light. The same comp nents, consequently, dictate the radiative pro erties of the planet's atmosphere. The altitud at which the infrared radiation leaves the Eart depends on the variable air humidity and carb nic acid content.

The equilibrium temperatures of other pl nets can be calculated in the same way as that the Earth. They depend mainly on the distan from the planets to the Sun. The farther the o bit, the lower the temperature on the planet surface.

The albedos of the solar system's bodies var within a wide range: from 6% of Mercury an the Moon to almost 100% of Enceladus, a sate lite of Saturn. The range of albedo alteration other large planets is not so wide. This metho being of little help in finding the accurate tempe ature conditions on each of them, let us ta the albedo of all the planets for our estimates equal to 40%. Thus calculated equilibrium tem peratures of the planets are cited in Table

Let us demonstrate now how the properties a planet's atmosphere could be estimated wi its temperature, radius, and mass being know You have seen on the Earth's example that th ability of a celestial body to retain the gaseou

Table 8

Atmospheres of the Solar System Bodies

Celestial body	Radius m	Temperature (K)	ξ_0	Description of atmosphere	Composition of atmosphere
Sun	7×10^8	5780	4000	The entire Sun is a gaseous sphere	e^-, H^+ H, He
Mercury	2.4×10^6	400	2.8	There is practically no atmosphere	—
Venus	6.1×10^6	290	22	Powerful atmosphere above solid land	CO_2, N_2
Earth	6.4×10^6	250	30	Powerful atmosphere above ocean and land	N_2, O_2, H_2O
Moon	1.7×10^6	250	1.4	No atmosphere	—
Mars	3.4×10^6	200	7.7	Weak atmosphere	CO_2, N_2
Jupiter	7.1×10^7	110	2000	Powerful atmosphere continuously passing into liquid	H_2, He, CH_4
Saturn	6.0×10^7	80	950	Powerful atmosphere continuously passing into liquid	H_2, He, CH_4
Titan	2.6×10^6	80	5.3	Powerful atmosphere above methane ocean (?)	N_2, CH_4
Uranus	2.5×10^7	55	500	Powerful atmosphere	H_2, He

Continu

Celestial body	Radius m	Temperature (K)	ξ_0	Description of atmosphere	Composition of atmosphere
Neptune	2.2×10^7	45	800	Powerful atmosphere	H_2, He
Pluto	1.4×10^6	40	1.7	No atmosphere (?)	—

atmosphere depends on the ratio of the therma velocity of molecules and the escape velocity Let us compile now another ratio which will als be a criterion for retaining the atmosphere. Com pare the energy of the gravitational attraction o a molecule with mass μ/N_A to its mean kineti energy. It is clear that if the ratio between thi gravitational energy $Gm\mu/N_AR$ and quantity kT would be low, the gas with molecular mas μ would fail to get into the atmosphere—it woul volatile very fast from the planet into space. I contrast to that, the ratio $\xi = Gm\mu/N_Ak$ should be much more than unity for a stable an powerful atmosphere to exist.

It can be easily seen that quantity ξ is a square ratio of the planet's orbital velocity to th thermal velocity of molecules. Let us calculat this ratio for the Earth at $\mu = 1$:

$$\xi_0 = \frac{Gm_\oplus}{kN_AR_\oplus T_\oplus} = 30.$$

The parameter ξ_0 of the Earth turned out to b notably more than unity which in fact corre

sponds to a large and stable atmosphere of this planet. Let us make a similar estimate for other celestial bodies. Calculate the ratio

$$\xi_0 = \frac{Gm}{kN_ART}$$

for the Sun, planets, and some of their satellites and compare this value to our knowledge of their atmospheres. In doing so, we shall consider the temperature of the solar surface as the temperature of the Sun and employ radiation temperatures of planets and satellites in order to estimate their temperatures.

Strictly speaking, in the case of Mercury, Venus, and the Moon it should be taken into account that they rotate slowly, therefore the temperatures of their illuminated and dark sides differ significantly and for retaining an atmosphere the temperature of the illuminated side is important. But the qualitative agreement has been reached all the same. Notice that the lesser ξ_0, the weaker the planet's atmosphere and the heavier its component-gas by molecular mass. At $\xi_0 < 3$ atmosphere is totally absent. Mars, with that value reaching 7.7, has a small atmosphere consisting mainly of heavy carbon dioxide. The content of water vapours having a low molecular mass is negligible in Mars's atmosphere.

Table 8 accommodates also Titan, a satellite of Saturn and the only, besides the planets, body of the solar system on which the atmosphere has been discovered. This fact was known from astronomical observations. It has been finally confirmed due to the flight of the American spacecraft "Voyager 1", which in 1979 and 1980 transmitted

to the Earth photos of Jupiter, Saturn, and the
satellites.

At the same time the large values of ξ_0 f
giant planets correspond to the fact that th
are composed of materials with low molecul
mass. Hydrogen and helium can be retained on
in atmospheres of sufficiently massive bodie
But the giant planets themselves have probab
become so giant just because hydrogen and hel
um are the most widespread elements in the sol
system. Apparently, the fall of the equilibriu
radiation temperature with the distance from th
Sun has resulted in the fact that the giant pla
ets consisting of light gases, light elements, a
located at the suburbs of the solar system, whi
the minor planets, the mass of which is built u
of a dense material, are located near th
Sun.

The explanation of the composition of atmo
pheres in terms of ξ_0 is attractive due to i
simplicity, but it fails to provide for an accura
estimation of the chemical composition of a
mospheres. It is also required that different g
ses of the atmosphere could coexist, that th
would not react with one another and with th
materials of the planet's surface. This conditio
however, is not absolutely necessary. For exam
ple, there is very much oxygen in the Earth's a
mosphere and this gas is very active and easi
enters into chemical reactions. There is quite
lot of combustibles and materials apt to oxidi
on the surface of our planet, recall, for instanc
wood fires. The oxygen equilibrium, as has be
already mentioned, is continuously maintain
by photosynthesis reactions. Thus, the atmo

pheres of planets should be stable in relation to the chemical equilibrium at constant irradiation of the half of the planet by solar light.

Apparently, carbonic acid was present in large quantities in the Earth's primary atmosphere which at that time was much more similar to that of Venus. For the present moment carbon, absorbed by vegetation in the course of photosynthesis, has been accumulated in geological deposits. These include coal, oil, natural gas, and chalk. Chalk has been formed from the residues of skeletons and shells of pre-historic organisms.

And how did the ocean come to existence? Are there oceans on other planets and what is their composition? First of all let us clearly formulate what an ocean is and under what conditions it can appear.

An ocean is a significant layer of liquid on the surface of a planet separated by a clear boundary from the atmosphere. Therefore temperature on at least a part of the planet's surface should be higher than the melting point of the material making up the ocean. This alone does not suffice however. If the total mass of this chemical compound on the planet is small, it could still be an atmospheric gas even if the temperature is all right. Consequently, the ocean's existence demands that the partial pressure of some gas near the planet's surface were higher than the saturated steam pressure.

Yet that is not all. Values of critical pressure and critical temperature are known for each compound. At a pressure above critical and a temperature higher than critical there is no physical

difference between a liquid and a gas, there is no boundary surface between these two phases and they continuously pass into one another. Hence follows that there could be no helium or hydrogen ocean on any planet of the solar system. The critical temperature of helium is merely 5 K, that of hydrogen is only 33 K. They are lower than radiation temperatures of all the planets. Helium and hydrogen abundant in the atmospheres of giant planets pass into dense layers of highly compressed material when they submerge into their interior in the same way as it happens on the Sun. Although this material looks like a liquid—distances between molecules approach the sizes of the molecules themselves—we would not call it an ocean since there is no boundary separating it from the atmosphere.

For the same reason there is neither methane nor ammoniac ocean on Jupiter, Saturn, Uranus, and Neptune although these gases are present in minor quantities in their atmospheres. At the altitudes, where the partial pressure of, say, methane would exceed the pressure of its saturated steam, it would start to condensate. But not into the ocean. Its total concentration being low, this condensate would be not a layer of liquid, not an ocean, but a layer of small drops of methane, a methane cloud.

To get a clearer picture of conditions for the existence of an ocean, assume for a moment that the temperature of the Earth's surface is higher than its actual level. Would the Earth's ocean boil if the temperature reached 100 °C? No, it would not. A part of water would evaporate,

pass into the atmosphere, so that the partial pressure of water vapours near the ocean's surface would amount to 1 atmosphere, which is the pressure of saturated water vapours at that temperature. The total atmospheric pressure on the Earth would be equal to 2 atmospheres; the other half of the pressure value will be provided by the partial pressures of nitrogen and oxygen. The ocean would still exist then. At which temperature, however, it would completely evaporate?

For this case let us first find the pressure at the surface of the dried-out Earth. If the entire Earth's water passed into the atmosphere, its pressure would exceed the present value by a factor of 267, since the mass of water is 266 times more than that of the atmosphere. However, this pressure of 267 atmospheres is higher than the critical pressure of water vapour equal to 218 atmospheres. Therefore the ocean's disappearance as the temperature increases would proceed as follows: the atmosphere would be gradually saturated by water vapour up to the moment when its pressure near the ocean's surface would reach the critical level. This would happen at the critical temperature which for water equals 374 °C or 647 K. Moreover, the boundary between the water and atmosphere disappears and the pressure at the former ocean's bottom would reach 267 atmospheres.

Let us turn back to the conventional terrestrial temperature equal to 15 °C on the average. At this temperature the pressure of the saturated water vapour equals 1700 Pa = 0.0168 atm. Hence it is easy to estimate the total mass of water

in the entire Earth's atmosphere:

$$m_{H_2O} \simeq 4\pi R_{\oplus}^2 p_{H_2O} \frac{1}{g} \simeq 9 \times 10^{16}\ \text{kg}.$$

It is clear that the mass of water on the planet must exceed this limit for an ocean to exist.

The only, besides the Earth, good candidate for possession of a real ocean in the solar system is Titan, a Saturn's satellite. In addition to nitrogen accounting for a major part of Titan's atmosphere, methane and hydrocyanic acid HCN were discovered there. The temperature on Titan's surface was measured by "Voyager 1" and appeared to be equal to 93 K which is higher than methane's melting point (90 K) and lower than its critical temperature equal to 191 K. The atmospheric pressure near the Titan's surface turned out to be 1.6 times higher than the pressure of the Earth's atmosphere. At temperature 93 K the pressure of saturated methane vapours equals 0.16 atm. Consequently, it would be sufficient for the existence of a methane ocean on Titan that the total content of methane in its upper layers were higher than 10%.

Assume that it is so, which is quite possible. And now imagine yourself the Titan's surface. There is a coast of methane ocean composed of the usual water ice with an admixture of frozen ammonia and hydrocyanic acid. A methane snow falls on grey-green methane waves. A dim light of the Sun—not brighter than the Moon light on Earth—hardly breaks through reddish methane clouds.

Then another vague and light circle sized as 12 Earth's moons ascends above the horizon. This

is the rise of Saturn*. One should immediately leave the ocean's coast and move the spacecraft to a highland, to some safe place. A tide will come soon and its height on Titan is assessed at about one hundred meters!

The ocean majestically upheaves, methane currents are seething, the splashes of surf wrapped up by vapours ascend to dark clouds completely hiding both the Sun and Saturn. The tide is over and a violent methane shower threatens to wash us down from faltering rocks...

3. The "Puff-Pastry" of the Earth's Atmosphere

The vertical layout of the Earth's atmosphere is very complicated and interesting. Density ρ, pressure p, and air temperature T all change simultaneously with height. It is easy to derive equations describing the dependences of pressure and density on height. One of them is the equation of ideal gas with molecular mass $\mu = 29$, which is the mean molecular mass of our nitrogen-oxygen atmosphere:

$$p = \frac{\rho N_A k T}{\mu}.$$

You are also well familiar with another equation: a hydrostatic equation describing the alteration of pressure in a liquid with height z.

* It is assumed here that Titan is not always turned to Saturn with one side, but rotates in relation to it. We do not know whether it is really so. The Titan's atmosphere is so dense that the details of its surface could not be distinguished so far.

However, the density of the atmosphere is varia ble in contrast to the density of an incompressibl liquid. Therefore, the hydrostatic equation shoul be written in a differential form for change i pressure dp with height increment dz:

$$dp = -\rho g dz.$$

The minus sign indicates that pressure decrease with height.

If density were excluded from these two equa tions, one could derive a differential equation o the dependence of pressure upon height:

$$\frac{1}{p} dp = -\frac{g\mu}{N_A kT} dz.$$

Note that the ratio of the increment in the poten tial energy of the molecule, $mgdz = \mu g dz/N_A$ to the specific kinetic energy of molecules, kT is placed at the right-hand side of the equation.

In case when the temperature does not alte with height, this differential equation corre sponds to a simple demonstration dependence o pressure on height. If the temperature of the at mosphere would remain equal to, say, 300 K, both the pressure and the density of the air wo uld increase twice with every 6 km of elevation.

In fact, however, the dependences of pressur and density on height notably deviate from such a demonstration dependence. This happens be cause the temperature of our atmosphere is far from being constant at different heights. It changes i a nontrivial way. Consider Fig. 31 and try t guess why this "puff-pastry" of cold and hot inter layers appears in our atmosphere.

Temperature decreases with elevation from the Earth's surface. At the height of 17 km in tropics and about 10 km in polar areas during the polar day it reaches the first minimum: —75 °C above tropics and —55 °C above the pole. This

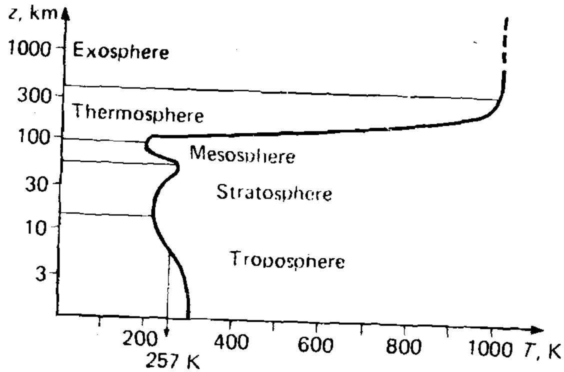


Fig. 31. Dependence of temperature on the altitude in the Earth's atmosphere.

is not a mistake, at such heights the temperature above the pole is higher than that above the tropics!

The area adjacent to the Earth's surface where the temperature falls is called troposphere. This word stems not from tropics but from the same Greek word *tropos* meaning "turn". The temperature turns above the troposphere while above the tropics the Sun turns during solstices.

The next layer of the atmosphere where the temperature increases is called *stratosphere*. This word originates from the Latin word *stratum*

meaning “pavement”. The minimum temperature between the troposphere and the stratosphere is called tropopause. The temperature rise in the stratosphere continues up to the height of approximately 55 km. At that height the temperature reaches its maximum called the stratopause and approaching 0 °C.

The place further upwards is taken by the mesosphere extending up to the altitude between 85 and 90 km. The word *meso* means “medium” in Greek. In the mesosphere the temperature falls again down to —85 °C. Above the mesosphere the temperature rises and reaches 1000 to 1200 K at the altitude of 400 km; this layer it termed thermosphere. Further upwards, in the external shell of the atmosphere called exosphere, the mean kinetic energy of molecules remains constant. It would not be quite correct to call it temperature: the density of gas is so low at those heights that molecules practically do not collide with one another. But specific energy of hydrogen and helium molecules at great heights is the same as at the temperature of 1000 to 1200 degrees.

“What a nonsense!” a mistrustful reader would say, “How can temperature reach 1000 degrees there when spacecraft fly exactly at those altitudes and cosmonauts are never boiled hard. If that were true, the sky there would have been red-hot.”

No, it would not. The air pressure at those altitudes is very low. At the altitude of 400 km it amounts merely to 10^{-8} mm of mercury column. The air density there is only 3×10^{-12} kg/m^3. In such atmosphere spacecraft meet a very weak resistance which allows them to stay in orbit

for years. The temperature of the atmosphere is high but because of the low density the environment does not affect the temperature inside the space stations. The “heat” outside does not disturb cosmonauts. The *internal temperature* of space stations is determined similarly to natural celestial bodies by the equilibrium of thermal energy carried away by the infrared radiation and the energy received from the Sun plus the own heat release.

But why the sky is not red-hot? This is a more complicated question. Have you ever noticed that a pinch of salt or soda placed into the transparent flame of a gas burner makes it bright-yellow and opaque. And be sure, its temperature remains the same. So the colour of a gas is determined not only by temperature but also by its composition and mainly by transparency. If the environment is transparent, its radiation is weak. But how could air fail to be transparent at high altitudes with its density being that low!

Now we are approaching the explanation of the puff-pastry puzzle. The key is concealed in the interaction between the solar radiation and various gases of the atmosphere since these gases absorb the light of different wavelengths. Figure 23 gives both the solar spectrum and Planck's spectrum of the Earth's radiation with temperature $T_\oplus = 257$ K. Those areas of spectrum are hatched which cannot directly reach a terrestrial observer because they are absorbed by atmospheric gases. It is partly a coincidence, but a lucky one for us, the residents of the Earth, that the major share of the solar radiation with wavelengths near the spectrum's maximum is not ab-

sorbed by the atmosphere and reaches the surface heating it thereby.

But where does the absorbed energy of the hatched spectrum areas go? You know already that it is retained by the atmosphere, but at what altitudes does it happen? Exactly at those at which Fig. 31 indicates a temperature higher than the radiation one: in the exosphere, in upper layers of the thermosphere and stratosphere.

Almost all gases making up the air contribute to the formation of the exosphere and thermosphere. Solar rays pass there through the first filter which cuts off the shortest wavelength area of the spectrum, the distant ultra-violet. This very part of the solar spectrum differs most from Planck's spectrum, the spectrum of the absolutely black body. We know already that this very part is related to the varying solar activity, the solar cycle. Therefore the temperature of the exosphere (it does not depend on the altitude) is significantly lower in the years of the quiet Sun (1000 K) and higher during the maximums of the solar activity (up to 1300 K).

As regards the stratosphere, it owes its existence to only one gas, ozone O_3. The area of its increased concentration is located at the altitude between 20 and 60 km but it completely absorbs the solar radiation within the range 220 nm $< \lambda <$ 290 nm already in the upper part of this layer. Although the spectral interval of ozone absorption is not wide, the share of solar radiation falling on it is 3 times more than in the entire, more distant, ultra-violet absorbed in the exosphere. However, air density is much higher in the stratosphere and this energy is dis-

tributed over a much greater mass of gas. This is why the maximum stratosphere temperature equal to about 0 °C is higher than the Earth's radiation temperature but significantly lower than the exosphere and thermosphere temperatures. There is no ozone above the stratosphere: the solar radiation breaks it into splits of O_2 and O.

The intervals between warm layers are taken by cold ones. Consider once again Fig. 31, the "puff-pastry" of the atmosphere. Both warm and cold layers emit energy into space by thermal radiation and exchange energy between themselves. However, the warm layers receive energy due to absorption of the solar light while the "edible" radiation fails to reach the cold layers and they get only "leftovers" due to mixing and over-radiation.

Certainly all this is only true of the day-side of the Earth. The night-side cools. Therefore, in all layers of the atmosphere temperatures are the lowest in the morning, before the sunrise, while the maximum of daily temperatures is, naturally, reached shortly before the sunset. Nevertheless, the layered structure of the atmosphere remains for the night; it lacks half a day to wash over, to destroy those five spheres: the tropo-, strato-, meso-, thermo-, and exosphere. Only the polar night features another composition of the atmosphere.

The first bottom layer of the atmospheric pastry is one more positive deviation from the Earth's equilibrium radiation temperature. The troposphere is the area of the maximum intensity of thermal processes and air motion, the area of the Earth's weather. The troposphere is heated

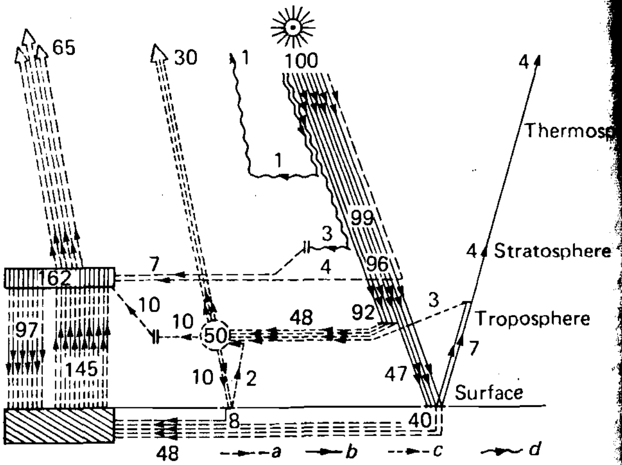


Fig. 32. Diagram of the Earth's thermal balance: *a*—infrared radiation; *b*—direct visible light; *c*—disperse visible light; *d*—ultraviolet radiation.

by the major part of the solar spectrum but i itself absorbs less than half of energy reachin it. The light is mainly absorbed by the Earth' surface, the land and ocean and then they hea the atmosphere from beneath.

Let us follow the redistribution of energ fluxes which the Sun sends to the Earth. Tak a hundred of conditional units of the solar powe which gets to the Earth and trace their paths i the atmosphere (Fig. 32). One per cent, the dis tant ultra-violet of the solar spectrum, is ab sorbed by molecules in the exosphere and ther mosphere heating them. Additional three pe cent of the near ultra-violet are absorbed b

stratospheric ozone. The infrared tail of the solar spectrum (4 per cent) remains in the upper layers of the troposphere containing water vapours (there is practically no water vapour above that level).

The remaining 92% of the solar light energy fall on the atmospheric “transparency pitch”: 290 nm $< \lambda <$ 2400 nm. This energy penetrates the dense near-surface air layers. A significant part of it (45 units) is dispersed by air mainly in the visible blue zone of the spectrum and gives blue colour to the sky. It is dispersed but not absorbed, it is not transformed into heat but changes direction: the straight solar rays turn into diffusive rays passing almost uniformly in all directions. The light is dispersed most intensively by clouds, the suspension of small water drops.

The direct solar rays—the remaining 47 per cent of the initial light flux—reach the surface. It reflects approximately 7 per cent of them and this light would share additional 3 units with the diffusive dispersed light of the sky on its way to space. Forty per cent of the energy of solar rays and eight units from the atmosphere are absorbed by the Earth's surface heating thereby the land and ocean.

The luminous power dispersed in the atmosphere (48 per cent taken together) is partly absorbed by it (10 per cent) and the rest is distributed between the Earth's surface and the space. Why does the space get more power (30 per cent goes upwards, 8 falls down) than the surface? The blue sky, however, seems to emit light uniformly, doesn't it?

This happens mainly due to the clouds. Th
are opaque and therefore throw back much m
light than they pass down. If you have ever flo
by plane, you must remember the dazzling wh
light of clouds illuminated by the Sun. Sm
clouds look bright from the Earth—they tra
mit enough light—but this light is much weak
than the eye-irritating light which they refle
upwards. Large storm-clouds are dark since a
most all the light falling on them goes upwar
after plural dissipation on water drops. In th
case the molecular absorption is very insignif
cant therefore, light coming from the clouds a
proaches the white spectrum, i.e. that of the s
lar light. Recall one more thing: when the pla
is landing it passes through a dense layer
clouds and a milk-white light surrounds you fro
all sides, but this light is brighter above than b
low. Every given volume of a cloud, within t
limits of which one can see, receives more lig
from above but disperses uniformly both upwar
and downwards. Therefore, the intensity of di
fusive light inside a cloud falls with descen
It is much darker at the bottom of a large clo
than at the top of it.

There is left 65% of the initial flux of the s
lar energy transformed into heat: 3 per cent
supplied by ozone, water vapours of the upp
troposphere account for 4%, 10% was absorb
in the main thickness of the atmosphere, and
nally, 48 per cent was transformed into heat
soil and in ocean water. A part of the latter en
gy returns to the atmosphere with condensati
of water vapours but this can be neglected f
the time being. All those 65 shares have been

nally transformed into heat and are carried into space by the thermal radiation.

Yet the entire spectrum of the Earth's thermal radiation—it is indicated in Fig. 23 together with the solar spectrum—falls on the area of absorption of water vapours and carbon dioxide. Thus, the luminous power which has been transformed near the Earth's surface into the thermal power cannot be liberated at once. It resembles very much the situation inside the Sun and a situation inside a cloud, only the luminous flux is infrared now, it is invisible to eye and propagates not downwards but upwards. In the same way as on the Sun, the radiation diffuses to outer layers until it reaches the altitude with so little absorbing gases (water vapours in this case) that they already cannot impede light or reradiate it.

This very altitude determines the radiation surface of this planet which is physically identical to the sphere which we call the surface of the Sun. It lies in the upper part of the troposphere, at the point of Fig. 31, where the temperature versus altitude curve crosses for the first time the straight line $T = T_{\oplus}$. Our planet would seem to have this radius for an observer looking at it through infrared glasses. In the diagram of the Earth's thermal balance (see Fig. 32) this process of radiation transfer is symbolically indicated by two meeting energy fluxes: 145 of our conditional units go upwards and 100 come down. This process of heating the Earth's surface and lower atmosphere is also called the greenhouse effect. It really looks like that. In a greenhouse solar rays also pass easily through a transparent

cover, heat the soil, while the heat cannot imm diately get away.

The figures of the thermal balance are averag for the entire Earth with no regard to geograp cal latitude. They almost do not depend on t season of the year. It is not easy to calculate th and the accuracy of their present estimation very low. The Earth's albedo itself taken here 35% is significantly lower according to rece measurements made from space (28%). Yet th thermal balance provides a clear explanation w our atmosphere resembles a temperature pu pastry and accounts for the mean annual tempe ature outside your door being notably high than the Earth's equilibrium radiation temper ture.

4. The Earth's Winds

Why does the wind blow? Put this question t your acquaintances and most likely they wi answer that the wind blows from an area wit higher pressure to an area of lower pressure. Th is partially true, but one thing is unclear: ho can pressure disturbances exist for a long time

Imagine a room with no draughts and no opera ing vacuum cleaner. The pressure would be co stant there at one altitude. In case of an art ficial pressure disturbance, for example, when tight balloon is pierced, you would hear a cla This is the result of air expansion. A further ex pansion due to inertia produces a rarefaction i the center which was filled again and so forth A sound wave has been generated; it reached th walls of the room weakening on the way, re

flected from them and returned to the center. The result is that after a very short time corresponding to room dimensions divided by the sound velocity the pressure disturbances would damp one another. In this case no directed air motion, the wind, would occur.

Imagine now one more thing: it is winter, there is a hot radiation under the window. The air rises above it, passes beneath the ceiling, descends at the opposite wall and having thus heated all the room moves over the floor back to the radiator while being cooled on the way. This is approximately the way the Earth is heated on the whole. Tropics function as a radiator, poles act as opposite walls and pressure deviations from the mean value are caused by temperature difference and, consequently, by the difference in air density. They are different in sign at different altitudes and cannot be equalized in few hours sufficient for the sound to travel round the entire globe.

You see that we are already familiar with the cause of transfer of air masses. This is convection, the lift of warm and light air replaced from below by cold air. The areas most heated during a day are the tropics, where solar rays fall on the Earth almost plumb. The temperature gradient, its drop with altitude near the surface, becomes more than the adiabatic, the equilibrium one, that is why a vertical air flux is generated. Air rises near the equator elevating the upper boundary of the troposphere. The altitude of the troposphere in tropics is 17 km, which is twice higher than its altitude near the poles. But where is this air to go? It is easy to understand that

at high altitudes it flows from the equator: t
northern air goes to the north and the southe
air rolls on to the south (see Fig. 33). Vertic
convection currents are turned into horizont
the warm air is partly cooled in the upper trop
sphere by sharing thermal radiation with spa

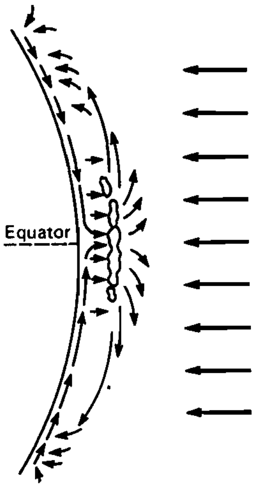


Fig. 33. Equatorial convec
tion—the cause of winds.

and descends in the middle latitudes and flow
back to the equator to compensate there the los
of air masses resulting from the convection lift

This is the operating cycle of the Earth'
heat engine. Its efficiency η is not high, merely
to 2 per cent. Let us estimate the velocity o
winds near the Earth's surface. We know th
power of the source of energy, this is $P(1 - A)$
the power of solar rays which get to the Eart
minus the power of the light reflected by th

planet. Its order of magnitude is 10^{17} W. In the final analysis this power is totally transformed into heat and then into thermal radiation but for a while a minor portion of it turns into mechanical power $P_{\text{mech}} = \eta P(1 - A) \sim 10^{15}$ W and operates the atmospheric heat engine.

The kinetic energy of terrestrial winds equals by the order of magnitude $m_A v_0^2$. We know the mass of the atmosphere and velocity v_0, the specific velocity of the Earth's wind, we are going to estimate. Winds born near the equator spread over the entire planet, although as we shall see, this motion is rather complicated. The characteristic time of the kinetic energy transfer equals by the order of magnitude $\tau \sim R_\oplus / v_0$, the time during which an air mass displaces over the distance of the Earth's radius.

To estimate the velocity of winds and the transfer time, let us put power P_{mech} equal to the value $m_A v_0^2/\tau = m_A v_0^3/R_\oplus$, the kinetic energy divided by the time of its transfer. Hence

$$v_0 \sim \left[\frac{\eta P(1-A) R_\oplus}{m_A}\right]^{1/3} \sim 10 \text{ m/s};$$

$$\tau \sim \frac{R_\oplus}{v_0} \sim 6 \times 10^5 \text{ s} \sim 1 \text{ week}.$$

Ten meters per second, 36 km/h, is it not too much for the average wind velocity near the Earth's surface? Firstly, this is merely an estimate and the other thing is that one should not judge this value by the experience of a resident of middle latitudes. The wind in the open sea is notably stronger than over the dry land and the velocity of ten meters per second is far from a hurricane.

The second estimate is very important. week is a characteristic time of weather alter tion. This is both the characteristic time of i fluctuations and the time being a physical boun ary between short-term weather alteratio caused only by motion of air masses and lon term variations effectuated by changes in th conditions of the heating of the Earth.

The estimated wind velocity v_0 is the velocit of air motion near the Earth's surface, since th mass of the atmosphere, the major part o which is concentrated at a low altitude, was in troduced into the kinetic energy. However, th air moves from the equator in the upper tropo sphere where its pressure and density are low It is clear that the velocity of wind at that leve should be sufficient to compensate its influ from beneath. Mass flows should be equal. A the altitude z this condition produces the follow ing estimate of the velocity: $v_z \sim v_0\rho_0/\rho_z$. I the upper troposphere the air density is by on order of magnitude lower than that near the surface. The velocity of wind there is higher b exactly the same value. In fact, winds blow a the altitudes of about 10 km with velocities approaching a hundred meters per second, hundreds of kilometers per hour.

Yet the real direction of wind there is not a all strictly to the north or to the south from th equator. The picture given by Fig. 33 is qualitatively true but only as a projection of wind directions onto a plane. The upper wind in the northern and southern hemispheres virtually deviates to become a west wind and the lower wind coming to the equator turns to the east direction.

This eastern wind prevailing in the open seas in tropical latitudes is called the trade wind. The cause of these deviations is the Earth's rotation.

To make it easier to understand how the Earth's rotation deflects the trajectories of motion, consider a clear example: the motion of a satellite. Let us have a look at a satellite's trajectory expressed in terms of terrestrial geographical coordinates. The satellite is acted upon only by the gravitational attraction and if its velocity is everywhere parallel to the Earth's surface, the curve of its motion will be a circumference of a large circle the center of which is positioned at the center of the Earth. The plane of this circumference is constant in relation to stars. Yet the Earth itself rotates and during the time of one satellite's turn T_s it rotates by the angle $\Delta\lambda = 2\pi T_s/P_\oplus$. Therefore, on the Earth's surface, in geographical coordinates, the satellite's trajectory looks like an oscillating unclosed curve. It can be seen on the photos of the electronic table operating in the Flight Control Center which demonstrates the motion of spacecraft over the Earth.

This trajectory is easily calculated. For this purpose, one should construct it on a motionless sphere, find the change in coordinates with time, longitudes λ and latitudes φ, and then displace the angle λ with angular velociy $\omega_\oplus = 2\pi/P_\oplus$, the velocity of the Earth's rotation. An example of satellite's trajectory is given in Fig. 34.

Newton's laws in their initial form hold true for inertial frames of reference, the systems moving uniformly and in a straight line. However,

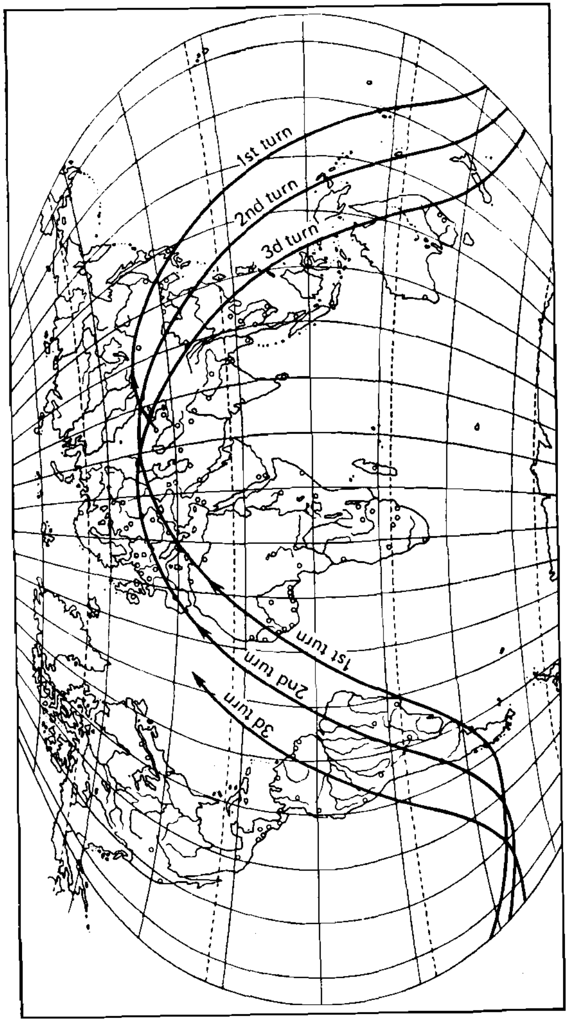


Fig. 34. Satellite's motion above the Earth's surface.

being positioned on the Earth, it is convenient for us to relate the frame of reference to its surface. This is the system of geographical coordinates. It is not an inertial system since the Earth rotates. The second Newton's law could be rewritten so that it held true for a rotating frame of reference. This, however, would require the introduction of an additional force into the right-hand side of the equation. This force is called the Coriolis force.

The Coriolis force is perpendicular to the velocity of the material point and proportional to the velocity. In addition to that, the Coriolis force depends upon the latitude of the location as sin φ. The general expression of it reads:

$$F_{\mathrm{Cor}} = 2m\omega_{\oplus}v \sin \varphi,$$

where m is the mass of the body.

Consider once again the satellite's trajectory. You see that there is no curvature at the points of crossing the equator and the Coriolis force is equal to zero at the equator. In the northern hemisphere, when the satellite moves eastwards, the Coriolis force is directed mainly to the south which results in the trajectory's return to the equator. In the southern hemisphere angle φ changes its sign and the Coriolis force turns the satellite to the north.

It should be noted once more that the Coriolis force is not a force caused by some real field of forces, it occurs only as the result of a description of phenomena in a rotating frame of reference. The fact is that the satellite's trajectory was first plotted in a motionless frame of reference, where it is just a circle, and the Coriolis force was em-

ployed only to explain it in the geographical coordinates.

However, one should not regard the Coriolis force as fiction since the effect of it is quite real for us. We are living on the rotating planet and the majority of phenomena which we observe occurs in the thin layer on the surface of the rotating sphere. A point moving over the sphere's surface passes from one radius of rotation to another. If we do not wish to notice it, if we want to regard the surrounding section of the Earth's surface as a motionless plane, *we should consider* the force of inertia, the Coriolis force, a real one. In the northern hemisphere it is directed to the right from the line of motion. The magnitude of the Coriolis force is insignificant. For example, in case of the velocity of a car equal to 20 m/s, the Coriolis acceleration reaches merely 10^{-3} m/s^2, which is only a hundredth of per cent of the free fall acceleration. Therefore, the Coriolis force practically does not affect our usual motions. *However, manifestations of this force become noticeable if the duration of its action is sufficiently long*: thus, the right-side rails of railroads have a shorter service life than the left-side ones, the right banks of rivers are steep while the left banks are gently sloping. Yet the Coriolis force has the most drastic effect on the global air flows.

Let us turn back to the tropical circulation. The velocities of air flows flowing from the equator to the upper troposphere are about 200 m/s. Naturally, besides the Coriolis force, the air masses are acted upon by other, aerodynamic forces. But look how the flows initially perpendicular to the equator change their direction under

the action of the Coriolis force taken alone (Fig. 35). *Firstly, the Coriolis force equals zero near the equator* and the flows maintain the same way as the trajectories of satellites, to the right in the northern hemisphere and to the left in the southern hemisphere. The turn by 90° occurs near la-

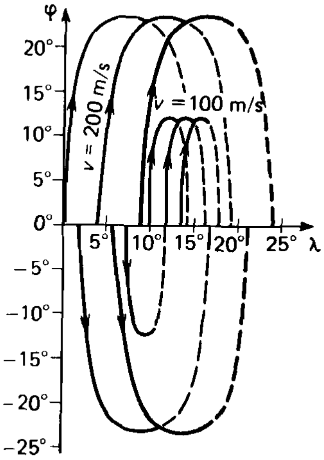


Fig. 35. Deflection of high-altitude winds coming from the equator by Coriolis forces.

titude φ_0 at which point air fluxes become directed from west to east in both hemispheres. This limiting latitude is related to the flux velocity as follows:

$$\tan \varphi_0 = \frac{v}{\omega_\oplus R_\oplus}.$$

For velocity 200 m/s this formula produces the latitude $\varphi_0 = 24°$. This however does not take into account the hydrodynamic, non-Coriolis, forces. In fact, this turn in the system of

tropical circulation of the Earth's atmosphe occurs approximately near latitude of 30°. A that point all the flows from the equator mer into two jet streams, the southern and the nort ern, two high-altitude winds blowing from we to east. They are called circumpolar vortice These vortices girdle the globe.

It should be noted that the jet streams shif to higher latitudes while crossing the oceans especially the Pacific Ocean. In contrast t that, they pass closer to the equator above th dry land. The cause of this phenomenon is quali tatively clear: the tropical convection is mor intensive above the oceans. Firstly, the ocean' albedo is lower than the dry land's albedo, there fore the absorbed solar energy is greater abov the open sea. Secondly, the air above oceans i humid. The density of humid air is lower tha that of the dry air at other similar condition since the molecular mass of water vapours i lower than the average molecular mass of air That is why the tropical convective stream rise higher above the ocean than above the dr land and impart a higher velocity to air masse leaving the equator.

The average altitude at which the jet stream are positioned is 9 km; at that altitude pressur equals 2×10^4 Pa which is one fifth of th pressure near the surface. The velocity in th middle of jet streams is 30 to 35 m/s, more tha hundred kilometers per hour. It takes a je stream 8 to 10 days, depending on the seaso of the year, to travel around the Earth. Soundin balloons filled with helium make sometime

several dozens turns round the Earth before they are carried out from the circumpolar vortex. This, however, concerns the southern hemisphere. As regards the northern hemisphere, two Americans made an attempt in 1981 to fly around the world on a balloon "Jules Verne" but failed. The crucial obstacle was the Himalaya, the highest mountain range of the world located exactly in the latitude of the jet stream.

The latitudes in which the circumpolar vortices are located depend, naturally, on the season of the year: in spring and autumn they are located almost symmetrically; in summer (in the northern hemisphere) they shift to the north—the northern vortex shifts to $\varphi = 50°$, the southern vortex moves over to 25° south. In winter it is vice versa: the northern jet stream descends closer to the equator while the southern travels to 45° south. The position of a circumpolar vortex does not remain constant also within one season. It often forms curves called meanders which slowly transit along the vortex. This phenomenon is one of the main causes of significant weather variations.

The convective lift carries air masses to high rarefied layers of the atmosphere. They expand and cool down. At the altitude where temperature falls below dew point, the air becomes saturated with water vapours which results in water condensation and formation of clouds. The Earth is constantly surrounded by a cloud belt near the equator. It is there that the convection, the vertical lift of the humid air, is most intensive and the sky above the equator is almost always covered by clouds.

Having risen to the altitude of 17 km above th tropics, air cools down to the temperature —75 °C and becomes very dry: it has left a most all moisture in the clouds at the altitud between 1 and 5 km. This air loses energy du to thermal radiation into space while descend ing and decelerating in the course of furthe motion from the equator and in jet streams. A masses of the convective flux accumulate im mense internal energy at the equator. The por tion of it spent on the elevation is released bac in the course of descent. Air passes the way from the equator to the latitudes of the circumpola vortex very quickly, approximately in a day During that time it cannot release much energ into space, therefore its temperature at th moment of descent back to the surface is high approximately 30 °C. The temperature at th equator is almost the same but the air is humid there, therefore its internal energy i greater.

The descent of air from circumpolar vortice to the surface occurs in latitudes between 2 and 30 degrees. The air is very dry and warm Have a look at the physical map of the world this very area accommodates the largest de serts of the Earth such as the Sahara in Africa the Arabian Desert and the Thar in Asia. Th southern hemisphere also features deserts a small areas of dry land located to the south of the circumpolar vortex: the Kalahari i Africa and several deserts in Australia. Th American continent houses less deserts (due t the Andes and Cordilleras) but the few it ha are located in the same latitudes.

The dry warm air descends from above and spreads over the surface. The velocity of its descent is not great and horizontal velocities are also insignificant. The latitudes under the circumpolar vortices are the areas of calms. The sailors have long ago named them the "horse latitudes". In the times of sailing vessels ships sometimes could not get out of there for months. Men suffered from heat and thirst but before them died horses transported by sea; a lot of horse skeletons lie on the bottom of the Pacific and Indian Oceans in horse latitudes.

The horse latitudes is an area of increased pressure and near the equator, in the zone of convective lift, the pressure is lower than normal. This refers to the standard measuring of pressure at the sea level. Why is this pressure difference practically constant, why is it not equalized with the velocity of sound as it happens in case of a hand-clap? However, let us consider the pressure distribution over height near the equator and under the stream jets. As we already know, pressure changes with height z as follows:

$$\frac{1}{p}\frac{dp}{dz} = -\frac{\mu g}{N_A kT}.$$

Temperatures of the surface near the equator and in horse latitudes are almost equal but the air near the equator is humid. Admixtures of water vapour reaching 3% by mass lower the mean molecular mass of gas since $\mu_{water} = 18$ is less than $\mu_{air} = 29$. Therefore the equatorial air is lighter, the right-hand part of the equation for it is less and, consequently, the pressure

drops with altitude slower near the surface tha
near the dry air.

The coldest place of the Earth's troposphere
positioned high above the equator; the temper
ture there is —75 °C, which is by 10 degre
lower than above jet streams at the same alt
tude. Since the humidity is negligible in bo
cases, the equatorial pressure drops with altitud

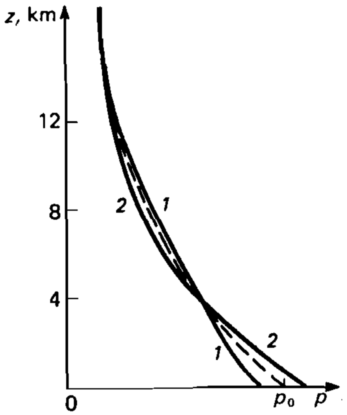


Fig. 36-37. Pressure
cyclone (*1*) and ant
cyclone (*2*) versus a
titude.

much more abruptly. Higher in the stratosphe
both dependences of pressure on altitude merg
The dependences of pressure on altitude plotte
for equatorial and horse latitudes look approx
mately as it is given in Fig. 36-37. You can s
that at the points where the pressure near th
surface is higher than normal it is lower in th
upper atmosphere than in an undisturbed atmo
phere at the same altitude. And vice vers
whenever we speak of a "decreased pressure" o
should bear in mind that it is lower than norm

only near the surface, where it is usually measured, while at the top it is higher than that of the average atmosphere!

Such pressure inhomogeneities produce not sound waves but air streams, the upper and the lower, directed at opposite sides. They are supported by air lift in the area of low (on the surface!) pressure and air descent in the high pressure area. The initial directions of these streams are altered by the Coriolis force clockwise in the northern hemisphere and counter-clockwise in the southern hemisphere. The lower streams directed from horse latitudes to the equator are trade winds, the north-east winds of the northern hemisphere and south-east winds of the southern hemisphere.

Almost the same explanation of trade winds was given as early as 1735 by the British scientist G. Hadley but, in his opinion, the atmospheric circulation was spreading from the equator to the poles. In his honour the tropical air circulation is called the Hadley cell.

And what are the average air directions in higher latitudes? The increased pressure near the surface in latitude 30° to 40°—under the jet streams—generates, besides the trade winds, winds directed at latitudes between 60° and 70°. At the same time the relatively low pressure above the jet streams draws air not only from the equator but also from high latitudes. Thus, another cell of atmospheric circulation in the opposite direction, the Ferrel* cell, is formed.

* A British school teacher who modified the Hadley system in 1856.

The circulation in the second cell is support not only by the high pressure in 30°-latitud but also by the low pressure near the ocea surface in latitudes $\varphi = \pm 60°$. It is caus similarly to the equatorial pressure, by conv tion. The energy for this convection is deliver by ocean currents.

Since the air motion in the Ferrel cell occu in the opposite direction compared to that the Hadley cell, the Coriolis force also deflec air streams in the opposite direction. For th reason west winds are prevailing in middle l titudes. In the southern hemisphere, where th only obstacle is the narrow band of the Ande this very west wind accelerates near the ocea surface almost to the speed of a hurricane. is not for nothing that sailors call these lat tudes the "roaring forties".

Finally, the air circulation resumes the straig direction near the poles. Air rises from the lo pressure areas and descends again near the pol producing there an increased pressure near th surface. East winds prevail near the poles pr ducing there an increased pressure near th surface. East winds prevail near the poles due the Coriolis deflection.

The diagram given in Fig. 38 describes on a very averaged system of terrestrial winds. So deviations from it are related to the dry la relief and the different albedo of the dry la and sea. The other deviations, the disturbanc changing with time, are, in fact, the proce which are used to call weather.

One more thing. One could get an impressi that the physics of the global circulation pr

osses is absolutely clear to us and we are just summarizing it here without the help of wise formulas. Alas, this is not true. The major difficulty is that climatic phenomena cannot be separated from weather phenomena. The fact is that winds cannot blow permanently as indi-

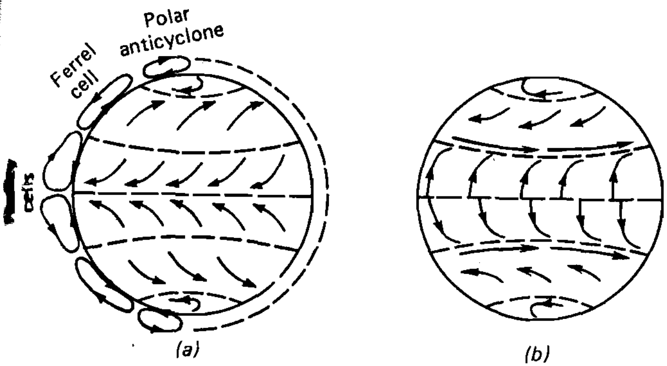


Fig. 38. Global pattern of terrestrial winds, cells of atmospheric circulation: (*a*) winds near the surface; (*b*) winds of the upper troposphere.

cated in Fig. 38! The variability, unsteadiness is an inalienable property of the Earth's atmosphere.

G. Hadley wrote in his work published in 1735 in which he defined trade winds almost in the same way as we do: "I think the causes of General Trade-winds have not been fully explained by any of those who have written on that subject". 232 years later the American scientist E. Lorenz, one of the modern authorities on the atmospheric circulation, began his book

with the following: "The opening words o
Hadley is classical paper afford an apt descriptio
of the state of the same subject today. Despi
many excellent studies performed since Hadley
time, no complete explanation of the gener
circulation of the atmosphere has been produced

5. The World Ocean and Its Currents

A part of this planet's surface, the Earth'
crust, is covered with ocean water. Ocean'
mass equals $m_{oc} = 1.37 \times 10^{21}$ kg. If the Earth'
crust were a regular sphere, the ocean's dept
would have been $m_{oc}/4\pi R_{\oplus}^2\rho_0 = 2750$ m at al
points. However, the shape of the Earth's soli
surface differs on the average by approximatel
the same value from an ideal sphere. Therefore
the ocean covers only two thirds of the Earth'
surface.

Consider how the Earth's firmament deviate
from the average sphere and the crust distri-
bution over the height of this deviation. Thi
dependence is called a hypsographical curve
Figure 39 indicates that high mountains an
deep depressions on ocean's bottom occup
merely 1% of the Earth's surface. The majo
area of the Earth is taken by two relativel
smooth surfaces. These are plains on dry lan
and ocean's bottom. There is quite a stee
altitude difference between them.

At first sight a hypsographic curve with on
gently sloping section, one bend, would be muc
more natural for a purely random surface. Thi
kind of curve, for example, is that of the wave
sea surface. Such is the surface of Venus studie

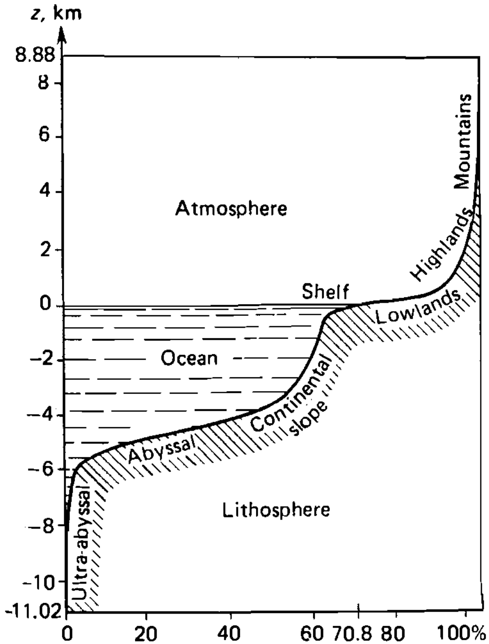


Fig. 39. Hypsographical curve—distribution of heights and depths of terrestrial crust.

by way of radiolocation through the dense and clouded atmosphere. The fact that the hypsographic curve plotted for the Earth's surface has two flat sections indicates a principal geological difference in the structure of continents, the elevated areas of dry land, and ocean bottom. The hypsographic curve is the result of the geological evolution of this planet.

This curve produces an averaged and thus highly smoothed profile of the transition from dry land to ocean's bottom. The typical dependence of the bottom surface slope on the distance from the coast is the following. The depth starts to increase slowly with the slope of merely 1.5 to 2 meters per kilometer. This low-water area around the continents extends from the coasts to the distance of 80 km on the average and is called the continental shelf.

Then comes a steep precipice and after 30 to 35 km, the depth reaches 3 km. Near the Ceylon coast the slope reaches even 30°. This area is called the continental slope. Then follows a small transition area succeeded by abyssal plains, a vast and almost flat ocean bed with depths between 3.7 and 6 km. The abyssal zone is crossed by very few submarine ridges and oceanic trenches, the depressions, the depth of which reaches 10 to 11 km.

The ocean resembles a water-filled vessel the edges of which reproduce the hypsographic curve. It is filled with water to the level higher than the average sphere of the Earth's crust by 2.44 km. The modern ocean covers 70.8% of the entire Earth's surface.

The quantity of water in the ocean is not strictly constant. It is known that the ocean level dropped more than once by 120 to 150 m. At that time shelf turned into dry land and the continental slope started immediately from the coast and at some places was even exposed. But where did those 150 meters of water get to? This water was accumulated on land in the form of immense glaciers, the ice mountains iden-

tical to those which presently cover the Antarctic and Greenland. During the glacial periods the surface occupied by the ocean decreased approximately by 5%. The ocean's area was always larger than that of the dry land.

The age of the Earth's ocean is not much less than the age of the Earth itself. Naturally, the ocean is in the state of almost complete chemical equilibrium with the Earth's atmosphere and its crust.

The ocean is a saturated solution of atmospheric gases. Yet their solubility in water is low. At 15 °C the share of nitrogen in the ocean's mass is 1.3×10^{-5}, oxygen accounts for 7.9×10^{-6}. The concentration of carbonic acid is more significant and important for organic life in the ocean: at that temperature it may reach 5.5×10^{-4} of the mass of sea water. It can be calculated that the total mass of carbonic acid dissolved in the ocean exceeds the total mass of atmospheric carbonic acid by 30 to 100 times.

During the history of its existence the ocean has dissolved the major part of chemical compounds soluble in water. These are mainly salts. It is more reasonable to specify the composition of ocean salt separately for cations and anions since the salts dissociate into ions in water (Table 9).

On the whole the ocean water is, naturally, neutral. The mass of anions is larger because the mass of a chlorine atom exceeds that of sodium.

The average content of this salt mixture in the ocean is 35.2 g per kilogram of sea water. This value, the sea water salinity is usually measured

Table

Composition of Ocean Salt

Cations	Share in mass	Anions	Share in mass
Na^+	30.60%	Cl^-	55.02%
Mg^{++}	3.68%	SO_4^{--}	7.71%
Ca^{++}	1.17%	Br^{--}	0.19%
K^+	1.13%	HCO_3^-	0.41%
Sr^{++}	0.02%	H_3BO_3	0.07%
Total:	36.6 %		63.4 %

by these units, i.e. thousandth parts. Anothe term for this dimensionless measuring unit i per mille* designated by $^0/_{00}$.

The sea water salinity changes significantl from place to place: vaporization increases sa linity while rains and river run-off dilute water As to proportions of salts indicated in Table 9 they are very stable. With the exclusion o the ion HCO_3^- whose concentration and th solubility of carbonic acid fall rapidly wit temperature rise, the salt content in seas an oceans practically does not depend either o geographical position or on depth. The constanc of soluble salts content was discovered by th first expedition on ocean study undertaken b the British ship "Challenger" between 1872-75.

The density of salt water is higher than tha of fresh water. At the temperature of 0 °C th

* The term stems from the Latin *pro mille* meanin "per thousand". Compare it to the word "per cent" (Lati *pro cent*, "per hundred").

density of medium-salinity sea water is 1.028 g/cm³, at 15 °C it is 1.026 g/cm³. Water density increases insignificantly with pressure. Even at the depth of 5 km where the pressure exceeds the atmospheric level by 500 times, the density of sea water at 0 °C is 1.051 g/cm³.

It is clear that the lower the depth, the higher the density of sea water. If this relationship is disturbed, a current occurs. This happens at the places where hot dry winds in a cloudless sky cause an intensive evaporation and salinity of surface waters.

A clear example of such a thermohaline* current is the water circulation in the Mediterranean Sea. A powerful stream of ocean water with salinity $36^0/_{00}$ flows into the Mediterranean Sea through the Straits of Gibraltar separating it from the Atlantic Ocean. The salinity of Mediterranean surface water increases to the east and near the Turkish coast it reaches 39 to $40^0/_{00}$. The salinizated heavy water sinks to the bottom heating thereby bottom waters. The depth of the Mediterranean Sea differs insignificantly from the average ocean depth but the temperature near its bottom is 12 °C which is by 9 to 10 degrees higher than the usual temperature of ocean depths. In the bottom layers the current flows in the opposite direction, to the west. Finally, the heavy salt water flows along the bottom of the strait back into the ocean. However, the total flow through the Gibraltar

* The Greek word ***hals*** means "salt". The word ***halogen*** meaning "producing salts"—an element of the 7th group of the periodic table—originates from the same root.

is directed towards the Mediterranean Sea since all rivers falling into the sea fail to compensate the loss of water due to evaporation.

It is interesting that the ocean water becomes less dense during the melting of glaciers, icebergs and sea ice, despite the fact that it is cooled thereby: desalination reduces density more than cooling increases it. Therefore, icebergs, blocks of ice splitted from the glaciers in the Antarctic and Greenland, float on beds of almost fresh light-weight water which mixes with the surrounding salt water rather slowly.

It may seem that the light water partially desalted and heated by the sun should flow over the surface from the Antarctic and Arctic to the low equatorial latitudes. This, however, does not happen and, apparently, did not happen when the Earth's climate was different. In fact the partially desalted surface water surrounding the Antarctic is mixed by the winds of the "roaring forties". In that area the cold Antarctic water, the salinity of which becomes almost equal to that of the underlying ocean water, sinks and reaches with its cold tongue the equator in the Indian Ocean and even the Tropic of Cancer in the Atlantic and the Pacific Oceans (see Fig. 40). The salinity reaches its minimum at the depth of about 1000 meters.

In tropics warm salty surface water gradually mixes with the cold partially desalted water of the Antarctic tongue. This is a rather peculiar process. The matter is that the temperature equalizes in differently heated water more than hundred times faster than salinity equalizes in water with nonuniform salinity. Thus, the pre-

sence of a layer of salty warm water above a layer of cold fresh water produces instability. This instability develops as indicated in Fig. 41.

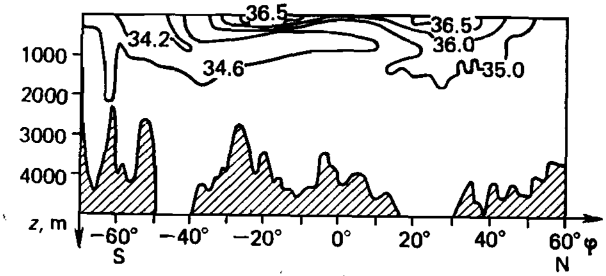


Fig. 40. A "tongue" of practically desalted ocean water from the Antarctic to tropics (cross-section in the Atlantic).

Assume that a slight bend of the boundary surface has occurred. The salty water in it cools faster than in the adjacent flat areas, but it

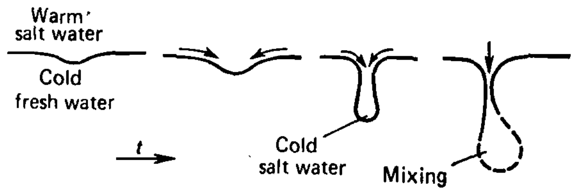


Fig. 41. Growth of salt fingers in foliated water.

fails to mix with the surrounding fresh water. The result is that this mass becomes denser than the surrounding liquid and starts to sink extracting with it a narrow stream of the upper salty

water, the so-called salt finger. Salt fingers split into more narrow streams and finally get mixed with the surrounding water deep below. Salt fingers are easily observed in laboratory conditions especially if salt water is coloured. The temperature and salinity difference is not high in the ocean but such an instability facilitates mixing.

The Antarctic and Arctic water is transported to equatorial latitudes deep under the surface, while the surface warm water is delivered from the tropics northward and southward. The major motive force for surface currents is wind.

Let us have a look how the wind generates current. Let a wind blow over a section of water surface with velocity W. It should be noted here that a strict calculation of the wind velocity is not an easy task. Firstly, this velocity increases with height, being practically equal to zero near the surface. Secondly, the wind above the sea is pulsating since it has to skirt the sea waves. Thus velocity W is assumed to be the average velocity of wind at the altitude of 15 m which is the height of masts of a small ship. At this level the velocity of wind starts to change slowly with height: at the height of 20 m the average velocity of wind is 1.037 W.

The friction of wind against water surface produces in it tension, tangential, and sliding shear forces. Shear stress σ is the frictional force per unit area. Therefore it is measured by $N/m^2 = kg/(m\ s^2)$. The value of friction being determined by the wind, it depends only on physical characteristics of air, its density ρ_1 and velocity W. These quantities can form only one com-

bination with the same dimensionality: $\rho_1 W^2$. Consequently, the shear stress is proportional to the density of wind's kinetic energy.

The experimentally measured dimensionless coefficient of proportionality turned out to be rather insignificant: 2.5×10^{-3}. The low coefficient is accounted for by the arbitrary selection of the altitude of wind measurement as 15 m. In fact, the viscous processes imparting tension to liquid occur in the air at a low level, where the air velocity is lower than W by an order of magnitude. Therefore, let us write down the dependence of shear stress on the wind velocity as

$$\sigma = \rho_1 (0.05\ W)^2,$$

which reflects the physical meaning of the velocity 0.05 W.

The shear stress causes the flowing of liquid with velocity v changing with depth z. In the course of sliding of upper layers against the lower layers this stress is balanced by the forces of viscosity

$$\sigma = -\eta \frac{dv}{dz},$$

where η is the coefficient of viscosity and the negative sign indicates that the viscous friction is directed against velocity.

Consider first a wind stream in a shallow but large water body. At the bottom of the water body, at depth D, the velocity of the current is equal to zero and the tension along the entire water level is constant. As a result the velocity

changes linearly with depth:

$$v = v_0\left(1 - \frac{z}{D}\right);$$

$$v_0 = \frac{\sigma D}{\eta} = \frac{D\rho_1}{\eta}(0.05W)^2.$$

Let us estimate the surface current velocit v_0 for a water body with depth, say, 1 m an air velocity $W = 1$ m/s. For this purpose w employ the table values of air density ρ_1 = 1.2 kg/m^3 and water viscosity $\eta = \eta_0 = 1.0$ × 10^{-3} kg/(m s). The result is absurd: v_0 = 2.9 m/s. But the velocity of a current cann exceed that of the wind which causes the curren Surely there must be an error.

The erroneous assumption was that the flo of liquid was treated as a plane current. Th would be true only for very low velocitie Calm and smooth liquid flows are called lamina flows. A laminar flow becomes unstable whe the velocity of current exceeds a certain limit. Th character of the flow of liquid or gas changes ab ruptly and the medium turns turbulent, th motion of individual points becomes chaotic Such a current is called turbulent flow*.

Formulas describing the viscous flow do no always hold true for turbulence. At best the can be applied only to the velocity average for small-scale turbulent pulsations. The turbu lent coefficient of viscosity, however, is alway significantly higher and may exceed the lamina η_0 by many orders of magnitudes.

* The corresponding Latin words are *lamina* meanin "a thin layer", "a plate" and *turbo* meaning "vortex" "perturbation"; there is also *turba* meaning "crowd".

Where is the boundary between laminar and turbulent flows? Beginning from what limit does the laminar flow become unsteady? In our example the characteristic velocity of liquid is v_0, the characteristic dimension of the problem is D. Consider a dimensionless ratio

$$\mathrm{Re} = \frac{v_0 D \rho_0}{\eta_0},$$

which is called Reynolds number. It turns out that a flow is laminar when Reynolds number is low and turbulent when this factor is high.

The boundary, the critical Reynolds number Re_{cr} depends on the real geometry of a flow. Its usual range is between 10 and 30. The turbulent viscosity grows approximately linearly with Reynolds number and passes into the usual laminar viscosity at $\mathrm{Re} < \mathrm{Re}_{cr}$:

$$\eta \simeq \eta_0 \frac{\mathrm{Re}}{\mathrm{Re}_{cr}} \text{ at } \mathrm{Re} > \mathrm{Re}_{cr}.$$

Let us again estimate the surface flow velocity under the same conditions but at turbulent viscosity $v_0 = \sigma D/\eta = \sigma \mathrm{Re}_{cr}/\rho_0 v_0$. Hence

$$v_0 \simeq \sqrt{\frac{\sigma\,\mathrm{Re}_{cr}}{\rho_0}} \simeq 0.05\,W \sqrt{\frac{\rho_1}{\rho_0}\mathrm{Re}_{cr}} \simeq 0.013\,W.$$

Thus, the flow velocity near the surface is proportional to the wind velocity. At the depth of 1 m and wind velocity of 1 m/s the order of magnitude of the surface velocity is $v_0 \sim 1$ cm/s and Reynolds number is $\mathrm{Re} \sim 10^4$.

This, however, cannot be applied to the real deep ocean. Between 1893 and 1896 a Norwegian explorer of the Arctic F. Nansen was drifting

with his ship "Fram" in the ice of the Arct Ocean and noted that at constant wind th drift occurred not in the wind direction but an angle of 20 to 40° to the right of it. Nans himself provided a qualitative explanation fo this phenomenon: in addition to the wind load the current is subject to the Coriolis acceleration Remember that it is caused by the Earth's rota tion with angular velocity $\omega = 7.3 \times 10^{-5}$ s^{-} and is directed perpendicular to the velocity, to the right in the northern hemisphere and to the left in the southern hemisphere. The magnitude of this value equals $2\omega v \sin \varphi$.

In 1905 the Swedish scientist W. Ekman devel oped a theory of wind-induced current in the open deep ocean. Look at the amazing turn o the current at depth due to Coriolis force (Fig. 42). Here are the formulas describing the Ekman spiral. If the wind is directed along the y axis, the current velocity vector $(v_x, v_y$ at depth z equals

$$v_x = \pm v_0 e^{-kz} \cos\left(\frac{\pi}{4} \mp kz\right);$$

$$v_y = v_0 e^{-kz} \sin\left(\frac{\pi}{4} \mp kz\right).$$

The lower sign refers to the southern hemisphere

The value of the surface velocity v_0 is approx imately the same as in the case of shallow wa ter but it is directed at an angle of 45° to the wind direction, to the right in the norther hemisphere and to the left in the southern hem isphere. The velocity vector turns with depth At the depth $z = 3\pi/4k$ it is reverse to the wind! The current velocity there equal

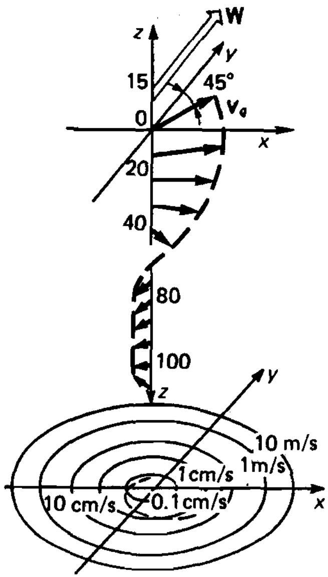


Fig. 42. Ekman spiral—wind-induced current of the deep ocean.

$v_0 \exp\{-3\pi/4\} = 0.095 v_0$. At depth π/k water flows in the direction reverse to the surface current with velocity $v_0 \exp\{-\pi\} = 0.043\ v_0$. Parameter k determines the characteristic depth at which the current turns and its velocity attenuates. It depends on the geographical latitude φ and is numerically equal to

$$k = \frac{\omega\,|\sin\varphi|}{v_0}.$$

At wind velocity $W = 10$ m/s the surface current velocity in the ocean is about 0.1 m/s and the depth at which the flow turns back is

about 100 m. At that wind velocity the Reynold number of the ocean turbulence is about 10 The turbulence mixes efficiently the ocean the surface layer down to the depth of 100 m The major transport of water masses by curren occurs in this very layer.

The advantage of Ekman's theory is th it gives a reasonable physical picture and a tr estimate of characteristic velocities and depths However, this is a very idealized design. It i inapplicable, in particular, near the equator where the Coriolis acceleration and paramete k turn into zero. But the main thing "preventing" the theoretical construction of the distributio of ocean currents is dry land, the continent of the Earth.

It is interesting to estimate the kinetic energ of all ocean currents. The average wind velocit on the Earth has been already estimated whe we calculated the kinetic energy of the atmosphe-re; this is $W \sim 10$ m/s, therefore the character-istic value of the velocity of ocean currents i $v_0 \sim 0.1$ m/s. The characteristic depth of cu rents equals v_0/ω. Without taking into accoun all numerical constants, multiply it by th area of oceans approximately equal to $R_\oplus^2$ an the density of the kinetic energy of currents $\rho_0 v_0^2$. Then the kinetic energy of the ocean wi be

$$\rho_0 \frac{v_0^3}{\omega} R_\oplus^2 \sim 10^{18} \text{ J},$$

which approaches quite closely the estimate o the total energy of individual currents. Thi energy is by three orders of magnitude les

than that of the atmosphere, which is natural, since the atmosphere is the ocean's major power supplier. Thermohaline currents fueled directly by the solar radiation are less intensive compared to wind currents.

Theoretical calculation of the real currents of the World Ocean is very complicated. One should take into account the inertia of currents deflected by coasts, the dependence of water temperature and salinity on depth, and the alteration of wind directions. Even powerful computers fail so far to produce a detailed map of ocean currents.

One more difference between the ocean currents and the idealized scheme is that they tend to form in the middle of the ocean narrow jets—merely 100 to 300 kilometers wide—flowing with velocity up to 2 m/s. The flow of such a river in the most powerful currents reaches 0.1 km^3/s. But an ocean river has no banks, therefore its position may change. The jet often makes a bend moving with the current. Such current bends are called meanders. The word originates from the ancient Greek name of a river in Asia Minor, the Maiandros. Modern maps name the river with a "Turkish accent": the Menderes.

The river flows in a very loose soil, erodes it and changes its bed very often. A meandering ocean current may fork, launch separate jets, make whirlpools several hundred kilometers in diameter. Such a spontaneous whirlpool moves slowly over the ocean and does not disappear for a long time, say, for about a year. The major stationary currents of the World Ocean also

have a habit of forming large whirlpools wi diameter of thousands kilometers.

Have a look at the map of main surface cu rents of the Earth's water shell (Fig. 43). Tra winds induce equatorial currents in tropics, latitudes ±15°, flowing westward. Taking in account that in these latitudes wind blows fro north-east in the northern hemisphere and fro south-east in the southern hemisphere, we se that the deflection of the surface current fro the direction of the wind really approaches 45 as dictated by Ekman's theory. In all the ocea there is an equatorial countercurrent directe eastward between the northern and the souther trade-wind currents. Its axis passes on th average about five degrees north of the equato there is *no complete symmetry* between th northern and the southern hemispheres of th Earth. There is also a reverse water flow in depth under the trade-wind currents. Being aware o Ekman's spiral, we are no longer surprised b its existence. The equatorial countercurrent ma be regarded as a surface outlet of Ekman's dee countercurrents at the places where wind is wea and Coriolis acceleration equals zero.

If liquid currents in Ekman spirals we summarized according to depth, it would tu out that the total flow is directed along the x-ax strictly perpendicular to the wind direction. Th total flow of each of trade-wind currents, su marized according to depths, is also directe not strictly to the west but has a compone directed from the equator. To compensate the off-flows, there is an elevation of deep wate at the equator. Deep water being cold, the te

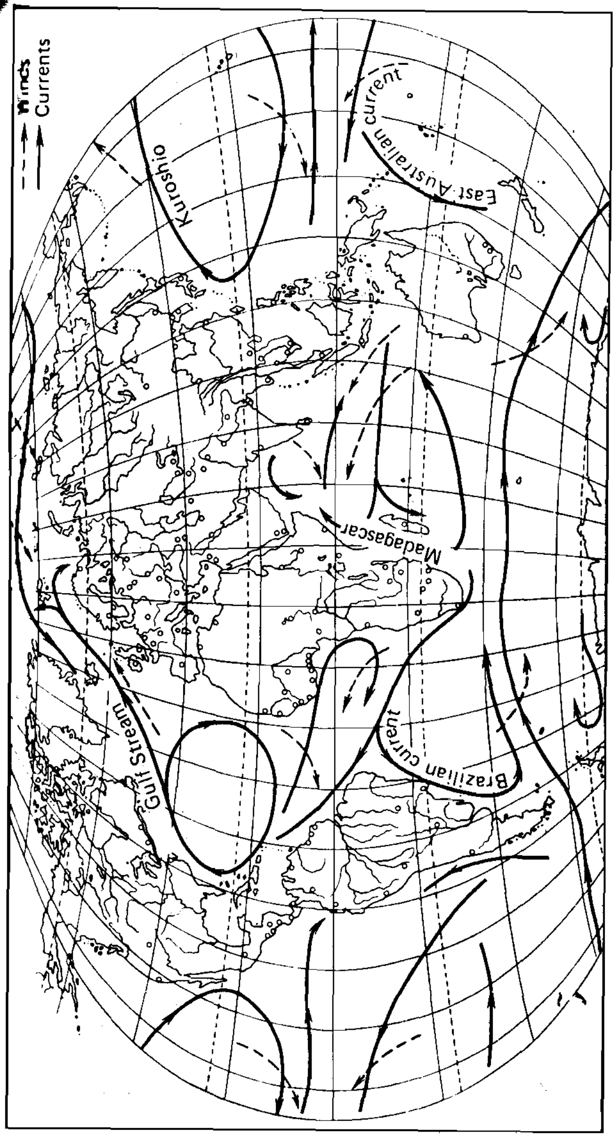


Fig. 43. Main currents of the World Ocean. Dashed lines indicate wind directions.

perature of surface water at the equator turns o to be by 2 to 3 degrees lower than the temper ture of the adjacent tropical water. The equato ial area of oceans is a relatively cold place of th planet! The slow elevation of ocean water

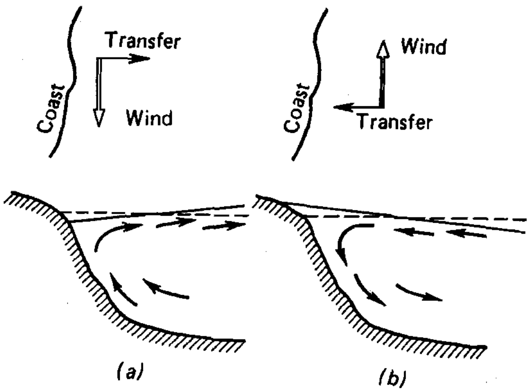


Fig. 44. Diagram of currents in the coastal area in th northern hemisphere; (*a*) upwelling—water elevatio (*b*) downwelling—water sinking.

termed "upwelling" in special literature; the term for sinking is "downwelling". In addition to th equatorial upwelling the elevation or sinking o water occurs at the coasts of large water bodie when the wind direction is tangential to th coast line. Such currents are exemplified i Fig. 44. This is also a manifestation of Corioli forces directing the total liquid flow, just as i Ekman's spiral, perpendicularly to the win direction.

When trade-wind currents approach the shel

they have to turn from the equator near the surface as well. This is how the powerful currents of middle latitudes start including the Gulf Stream*, and the Brazilian currents in the Atlantic, the Curosio, the East-Australian currents in the Pacific, and the Madagascar stream in the Indian Ocean. The major motive force of these currents is not the wind but the water head near the continents. The water level in the Mexican Gulf of the Atlantic Ocean is by 60 cm higher than that at the African coast, while the difference between the levels of the east and west coasts of the Pacific Ocean reaches 70 cm. This surge is induced by trade winds. Yet the structure of currents in middle latitudes is not wind-induced; they are traced in the ocean down to the depth of 1.5 to 2 kilometers and a weak countercurrent can be located only at still lower depths.

However these currents, especially in latitudes between 40 and 50°, are also pushed and increased by winds. Continents and the Coriolis forces deflect them to north-west in 30°-latitudes and in latitudes between 40 and 50° they flow straight westward. At that moment they are again—have one more look at the map of winds—reinforced by winds blowing from the south-east in the northern hemisphere and from the north-east in the southern hemisphere. The most powerful of the currents thus reinforced is

* The name stems from the Mexican Gulf where the current starts. The Japanese word *Curo shiva* means "a black flow". The colour of water in the Gulf Stream and Curosio is dark-blue, different from the colour of surrounding water.

the westwards Antarctic circumpolar curre which collects the middle-latitude currents all the oceans and which does not have cont nental obstacles on its path and is rushed north-west winds.

Finally, there is an almost circular West Arct stream in the Arctic Ocean induced by north east winds of the third, polar Hadley cell. I the southern hemisphere it has a weak analogu near the very Antarctic coast.

This is the global layout of ocean current logically connected with the map of winds an the map of continents. There is however one mor trouble. We have seen that both the winds o the Earth and the currents of the ocean ar significantly deflected by the planet's rotation, the Coriolis forces. The question is: what is the influence of the winds and currents on the Earth's rotation, don't they brake it?

The question seems very complicated. But the answer is very simple: no, they do not. Both the winds and currents are induced by the solar radiation falling onto the Earth. It supplies power to the motion of the atmosphere and ocean. This power is dissipated and transformed into heat. This really generates frictional forces but these forces are internal. According to Newton's third law, for every force retarding the planet's rotation there is an equal reverse force accelerating the rotation of the Earth. The total moment of all internal forces equals zero. The solar radiation does not change the Earth's moment of momentum winds and currents neither slow down nor accelerate the Earth on the average.

Chapter 5

The Earth's Climate

1. Clouds

We have already discussed the significant contribution of clouds to the Earth's albedo, its mean coefficient of reflection of solar rays. Yet it is not the only influence of clouds upon the Earth's climate. We all know how much weather depends on clouds. If the wind failed to transport clouds, the dry land, the continents, would become almost waterless. The major part of precipitations comes from the clouds: rain, snow, and hail. Finally, the role of clouds in the planet's thermal balance is not a purely negative one. They cover the Earth like a blanket, especially at night, decreasing thereby its cooling by radiation. But what is a cloud in fact?

The famous Russian dictionary by V. Dal' reads: "A cloud is a fog in the height". It is true. The fog, which everyone has seen and touched, is really a suspension of minute and minutest droplets.

A cloud consists of water drops and water is heavier than air. Why doesn't the cloud fall, why doesn't it sink completely? What supports it high in the air? And there is one more question. When a car moves in the fog the driver has to turn on the windscreen wiper. Drops of fog colliding with one another, grow slowly, fall down, and precipitate. The same thing

should happen in a cloud. Why does it n rain then from every cloud?

Let us first consider the motion of an indivi ual water drop in the air. A small drop is sphe ical as a result of water surface tension: a sphe gives a minimum surface area for a given vo ume. Let the radius of the drop be R, then it volume will be $V = 4\pi R^3/3$ and mass will b $\rho_0 V$. The weight of the drop $\rho_0 gV$, is man times more than the Archimedean force $\rho_1 g$ since air density ρ_1 is by far lower than wate density ρ_0. Thus, at first sight it may seem tha the drop must fall down with the acceleratio almost equal to g.

But as soon as the drop starts to move, a force resistance, the viscous force, occurs. For spheric particles this force is also called the Stokes forc after the British physicist J. Stokes who solve in 1851 the problem of the motion of a viscou medium around a sphere. The Stokes forc balances gravity, therefore the drop falls dow not with a constant acceleration but with constant velocity.

The Stokes force is a force of resistance pro portional to air viscosity $\eta = 1.8 \times 10^{-5}$ kg/(m s) The dimension of this force can be derived fro the product of the dimension of viscosity b the dimension of the velocity of fall $[v] =$ m/ and by the dimension of the drop's radius $[R]$ = m. In fact, the Stokes force, which certainl cannot be found with the help of the method o dimensions, equals the product $\eta R v$ with a accuracy to a numerical coefficient 6π:

$$F_S = 6\pi\eta R v.$$

Equating it to the drop's weight $\rho_0 gV = 4\pi R^3\rho_0 g/3$, it is easy to find the dependence of the velocity of fall on the radius of the drop:

$$v = \frac{2}{9}\frac{\rho_0 g R^2}{\eta}.$$

Let us verify now for which drop radii the air motion around the falling drop will be laminar. The formula of resistance was in fact derived by Stokes for a laminar flow. While treating the mechanism of ocean currents we have found that it is possible to determine whether a current is laminar or turbulent by way of calculating the dimensionless Reynolds number $\mathrm{Re} = Rv\rho_1/\eta$. The Reynolds number includes air density ρ_1, instead of water density ρ_0, because it is the air and not the water in the drop that is moving, flowing around the drop.

If the Reynolds number does not exceed several tens, the current is laminar, the Stokes formula is true and, consequently, the velocity of the drop's fall is really that. Let us introduce the velocity into the expression of the Reynolds number, equate it to 10, which is approximately the critical Reynolds number for the flow around a spherical body, and find the maximum radius R_1 at which the air flow around a water drop is still laminar:

$$R_1 \simeq \left(\frac{45\eta^2}{\rho_0\rho_1 g}\right)^{1/3} \simeq 0.01 \text{ cm}.$$

Such a drop with a radius of 100 micrometers falls with velocity $v = 1.2$ m/s.

If the radius of the drop exceeds the critical, R_1, the flow around the drop becomes turbulent

and the force of air resistance is no longer ex- pressed by the Stokes formula. It can be estim- ated by the order of magnitude assuming, as i the ocean, that the turbulent viscosity is pro portional to the Reynolds number. In that cas $F_{res} \sim \rho_1 v^2 R^2$ and equating this force to th weight of the drop produces the velocity of fal of the drop with radii larger than R_1:

$$v \sim \sqrt{\frac{g\rho_0 R}{\rho_1}}.$$

Note that velocity grows with the radius of th drop considerably slower than in the case of laminar flow.

Finally, new effects come into play with further increase in the drop's radius. It turn out that the surface tension of water $\sigma_0 = 7.2 \times 10^{-2}$ kg/s² fails to maintain the spherica shape of a large drop. Compare the force o turbulent resistance F_{res} to the force of th surface tension $\sigma_0 R$. They are similar by th order of magnitude when the drop's radiu equals

$$R_2 \sim \sqrt{\frac{\sigma_0}{\rho_0 g}} \sim 0.3 \text{ cm}.$$

Drops of that radius are flattened by the on rushing air flow. Then, with the radius increa ing the drop takes an irregular form changin in the course of falling and finally, when th radius of several R_2 is reached, air splits th drop into two parts.

Thus, drops break apart when drop rad exceed 0.3 cm; at 0.01 cm $< R <$ 0.3 cm, the fall down with the velocity proportional to $\sqrt{\ }$

and at $R < 0.01$ the velocity of fall is proportional to R^2.

This however, is not a complete picture. The smallest droplets of a cloud behave differently. They do not fall or, more precisely, almost do not fall. The size of particles being fractions of a micrometer (light clouds are composed of just such particles) the fall in the field of gravity can be neglected compared to an intensive Brownian movement*.

Imagine a very small droplet the size of which, however, by far exceeds the molecular size. The surrounding air molecules with masses $m \sim \mu/N_A \sim 5 \times 10^{-26}$ kg move with the velocities of about $v \sim \sqrt{kT/m} \sim 300$ m/s. The droplet moves chaotically under their impacts.

The mean kinetic energy of air molecules at temperature T equals by the order of magnitude kT. As a result of numerous collisions the molecules exchange energy between themselves and impart energy to the drop. This occurs in conformity with the principle of equipartition and the mean kinetic energy of the drop also approaches kT. Hence it is easy to estimate the root-mean-square velocity of the Brownian movement of the drop with the mass of about $\rho_0 R^3$:

$$v \sim \sqrt{\frac{kT}{\rho_0 R^3}}.$$

The Brownian velocity depends on the radius of the drop as $R^{-3/2}$ and the velocity of fall

---

* It was discovered in 1828 by the British botanist R. Brown.

of small droplets grows as R^2. These velociti
reach a similar order of magnitude at the rad
of the drops approaching

$$R_0 \sim \left(\frac{kT\eta^2}{\rho_0^3 g^2}\right)^{1/7} \sim 1.5\times10^{-4} \text{ cm}.$$

Droplets with radii less than several micro
meters, the Brownian particles, are movin

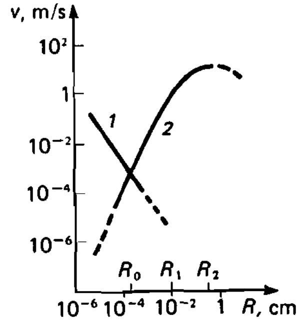


Fig. 45. Velocity of motion of water drops versus radiu
1—Brownian movement; *2*—fall.

fast, chaotically and to all directions. Dro
with radii larger than R_0 sink down. The slowe
drops have radii about a micrometer and mo
at a speed of about 0.03 cm/s. Velocities of wat
drops in air are indicated in Fig. 45 versus t
full range of their radii.

A real cloud accommodates drops of differe
dimensions. They move, collide, merge, a
grow. Brownian droplets grow rather slowly
the course of collisions. Their formation, grow
or evaporation is determined not as much by t

frequency of collisions as by air humidity and temperature.

For every liquid there is a dependence of the concentration of its saturated vapour upon temperature. Figure 46 gives such a dependence for

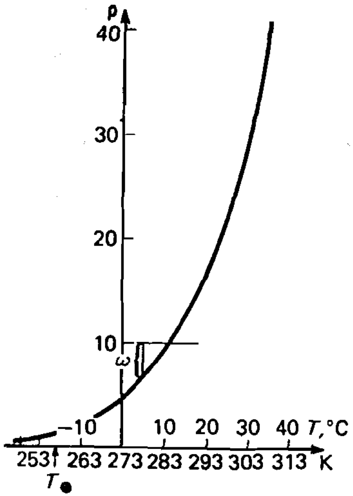


Fig. 46. Density (ρ, g/m³) of saturated water steam. Alteration of air humidity and water content.

water vapour. If the temperature of humid air containing, say, 10 g of moisture per cubic meter drops lower than 11 °C, the water vapour becomes saturated. A further cooling of this air mass will leave as much moisture as allowed by the curve in Fig. 46 and all the remaining water will be condensed.

This very condensate of the minutest water droplets in the air gives birth to fogs and clouds.

The mass of water drops and frozen water in unit of air volume is called water content of cloud, w. The water content of white clou is low while that of thunderstorm clouds reach 3 to 5 g/m^3.

But then, a drop has exceeded the radius R and started to sink. It grows now due to co lisions with other small droplets and the growt of size is accompanied by an increase in th velocity of fall and cross-section of drople capture. During the time dt the drop will trave the distance $dz = vdt$ and sweep the volum $\pi R^2 vdt$. Assume that it will collect all mino drops in this volume. Multiply it by the cloud' water content w to find the increase in mass o the drop:

$$dm = 4\pi\rho_0 R^2\, dR = w\pi R^2\, dz = \pi R^2 wv\, dt$$

$$= \frac{2\pi}{9}\, w\, \frac{\rho_0 g R^4}{\eta}\, dt.$$

The last equation employs the Stokes veloc ity of drop's fall. The second and the fifth expres sions of this chain of equations produce th equation for the drop's growth, the equatio of the time-dependence for its radius. Henc it is easy to see that almost all the time of drop' ripening, τ, from the small, Brownian radii R_0 to the large radii $R \geqslant R_0$ is taken by th initial period of growth. The water conten being constant along the drop's way. the ri pening time will be

$$\tau = \frac{18\eta}{wgR_0} \simeq \frac{6 \text{ hours}}{w\,[\text{g/m}^3]}.$$

It does not depend on the drop's final radius! However, a life 6 hours long is rather short for a cloud with water content of about 1 g/m^3. Normally, it never rains from such clouds. What is the reason?

Consider the second and the third expressions of the same chain of equations. They indicate the increment in drop's dimension with sinking by dz. The drop's radius grows independently of its velocity, with no regard to whether its fall is laminar or turbulent:

$$\frac{dR}{dz} = \frac{w}{4\rho_0}.$$

Therefore, a large raindrop can grow up to a radius approaching R_2 only in a powerful cloud with high water content. Even if the mean water content of the cloud is equal to, say, 3 g/m^3, the thickness of the cloud should not be less than

$$\Delta z = \frac{4\rho_0}{w}\, R_2 = 4 \text{ km!}$$

If the cloud's thickness is insufficient, the final size of drops would be so small that they would evaporate before reaching the Earth. This is the answer to the question, why not every cloud produces a rain. Special conditions are required for a many-kilometer-high cloud to form.

But what conditions exactly? What determines in general the upper and lower boundaries of a cloud?

Recall the "puff-pastry" of the atmosphere. Water vapours do not pass the infrared thermal radiation of the Earth's surface. Therefore ther-

mal energy is irradiated into space by the up layers of the troposphere where the content water vapour is insignificant since water been almost completely condensed and fro lower in clouds. At the upper boundary of clo the temperature is approximately equal to Earth's radiation temperature, 257 K = — 16 and saturated water vapour (its density at t temperature equals only 1.27 g/m³) becom transparent to the thermal radiation. Besid that the density of air and the density of wa vapour drop quickly with elevation. Theref the upper boundary of clouds is determined the total moisture content in the air and go near the level at which the infrared absorpti of saturated water vapour becomes so low that t thermal radiation can escape to space.

On the other hand, while descending in t cloud temperature increases. Concentration saturated vapour increases with temperatu very fast. Thus, at a certain altitude temper ture becomes sufficient for evaporation of water drops of the cloud and its water conte turns into zero. By this reason the lower boun ary of the cloud is determined by the dew poi the point in Fig. 46 at which the moisture co centration line in the cloud's lower part cros the curve of the saturated vapour concentratio

Clouds fall into two main categories. T first is termed stratus. The strati are formed the course of cooling the low-mobile air mass It usually occurs at night when the clou upper boundary releases thermal radiation in space. The other cause of their formation is t motion of warm humid air mass above the co

earth surface or above a cold air mass. A fog is just a stratus which lower boundary passes immediately near the Earth or the sea surface. The rain, if any, from such a cloud is a drizzle, light rain, the stratus lacks thickness to produce large drops.

The other type of clouds is called cumulus. The cumuli is the result of convection of the air rich with moisture. In the course of convective elevation air cools down adiabatically and, at a certain altitude, its humidity becomes saturated. This is the lower boundary of a cumulus. This boundary itself is almost motionless but air constantly passes through it. Air elevates also in the cloud itself while increasingly cooling down and condensing its moisture. At the cloud's upper boundary the cooled air, leaving all its water in the cloud, flows aside and sinks around the cumulus. This is why the cumuli often look like white lambs surrounded by cloudless intervals. If you have ever seen a large cumulus from above, from an aircraft, you could have happened to see the impressing view of convection cells located regularly either in even rows or individual hills rising in chess-board order. The upper part of cumuli usually consists not of drops but of ice crystals.

A powerful convection gives birth to a storm cloud, storm cumulus. Its usual height is between 7 and 10 km and very rarely (but always near the equator) it may reach 12 to 15 km. The structure of a storm cloud is more complicated, it houses both upward and downward air streams. Air is pulled downward by falling pieces of ice and raindrops.

The classification of clouds, however, is n completed by these two types, the strati a the cumuli. There are mixed types of clou and one more independent type, the cirrus whi means "curl" or "lock" in Latin. The cirri co sist of fine ice crystals and are formed at hi altitude in fast turbulent wind jets.

There are also clouds on all other planets wi powerful atmospheres. They cover complete the surfaces of Venus and Titan. As to the su faces of Jupiter and Saturn, they are just ide tified with the upper boundary of clouds sin they do not have any other solid or liquid su face. The chemical composition of clouds o other planets is very diverse. For example, som clouds of Venus are drops of sulphuric acid!

It can be assumed that mechanisms of clou formation on other planets are similar to tho on the Earth. However, one should bear mind that the formation of clouds can be caus not only by the condensation of drops and cry tals but also by a chemical reaction. A labor tory example of such a possibility is the fo mation of a small cloud of crystals of ammon um chloride NH_4Cl above bottles with hydr chloric acid HCl and ammonia solution N placed near to one another.

The clouds on the Earth are a significant fe ture and reflection of the Earth's weather. Nor ally, heavy clouds are arranged above the plac with low surface pressure. The surface win rush to such places being twisted by the Corio forces (Fig. 47). In the center of such a cyclon

* The word originates from the Greek word *kyklō* meaning "wheel", "coil".

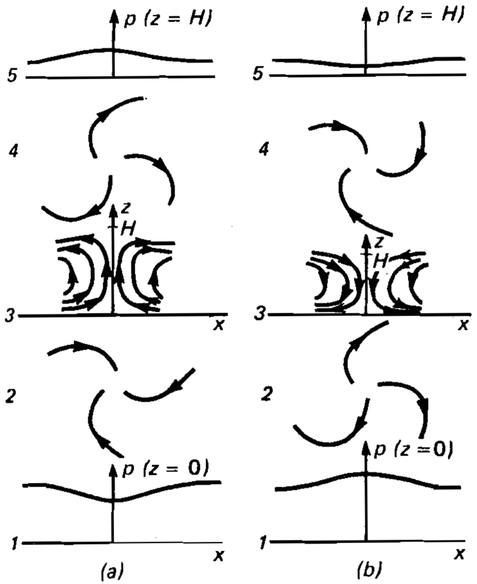


Fig. 47. Diagrams of cyclone (*a*) and anticyclone (*b*)
1—pressure near the surface; *2*—directions of surface winds; *3*—vertical cross-section; *4*—directions of high-altitude winds; *5*—pressure of the upper troposphere.

air rises upward, cools down and forms clouds. In the upper layers of the cyclone's atmosphere—above the low pressure area—the pressure of atmospheric air is higher (!) than the mean value, characteristic of this altitude. In the upper troposphere air pushed by the excessive pressure flows from the center of the cyclone.

An area of increased atmospheric pressure near the surface is called anticyclone. Dry air sinks

from the upper troposphere in an anticyclone. This is why the sky above the anticyclone area is generally cloudless and clear.

Cyclones and anticyclones are 200 to 3000 km wide and live for a weak on the average. However, some of them are located at one and the same place for a very long time. There is one permanent cyclone on the Earth which stands in summer and in winter near Iceland. It is generated from the meeting of the warm Gulf Stream water with cold polar air. Cloudy skies cover Iceland. In winter the weather in the USSR is mainly determined by the Siberian anticyclone. The major contributor to its formation are the Himalayas which do not pass the moist air of the Indian Ocean to the north. Yet the majority of pressure anomalies does travel—in general it coincides with the directions of global winds of the Earth indicated in Fig. 38*a*.

The number of cyclones and anticyclones on the Earth is approximately the same at any moment of time. Clouds cover about half the surface of this planet.

2. What is Weather and What is Climate?

We know very well from everyday experience what is weather and what is climate. What is the weather today? You look out of the window to see whether it is raining or snowing, look at the thermometer, listen to a weather forecast on the radio and put on the appropriate clothes. Climate is also a clear thing: it is cold in the north and it is hot in tropics throughout the

year. Yet we shall have to make these intuitive ideas a bit more accurate.

Weather is the state of the atmosphere and the Earth's surface (dry land, ocean) at the given day and hour on the entire planet. Usually we limit weather descriptions to individual areas of the Earth and characterize it by meteorological parameters. These include temperature, pressure, and air humidity near the surface, velocity and direction of wind, cloud conditions, the quantity and nature of precipitation. The modern idea of the Earth's weather is the totality of such data collected by all weather stations plus photographs of clouds made from satellites.

Unfortunately, this information is not complete. One of the reasons is that the number of weather stations is insufficient so far, especially in oceans. Another reason is that the complete information on weather is not several values measured near the surface, but their dependences through the entire tropospheric column, the lower layer of the atmosphere. From the theoretical point of view it is sufficient to have two altitude dependences, for example, for temperature and for humidity. These dependences being known, all other physical values can be found: air pressure and density at any altitude can be calculated, the upper and lower boundaries of clouds can be determined, water content of clouds can be estimated which, consequently, allows us to find out whether it's raining. The distribution of pressures makes it possible in principle to calculate winds at all altitudes; winds and heating of the ocean determine the

ocean currents. Finally, the tomorrow weathe depends on the motion of air and water masse

Certainly, we do not have such a complet information in real life and, probably, will neve have it. But even if the information on weath *conditions*, the height dependences of thes two functions over the entire planet, were com prehensive and sufficiently accurate, we coul not deal with a volume of information tha much: our computers would calculate weathe much slower than it changes in nature. However let us assume that a computer capable of cal culating with any required speed has been de veloped. Could the weather be forecast with an degree of accuracy and for any term in the fu ture?

No, it could not. The accuracy of a forecas would be very high for one day, satisfactory fo one week, a one-month forecast would be rathe approximate, and estimates of a one-year forecas would deviate from real weather conditions b the order of magnitude approaching deviation from the average many-year climatic data. Thi does not mean that long-term forecast is complete ly impossible but efforts in this direction hav a certain limit. At any accuracy of calculation divergence would grow as a square root of tim

This may seem surprizing since we have a sumed that the information on weather w comprehensive and the accuracy of calculatio was adequately accurate. Yet this is a propert of hydrodynamic equations. From the momen a motion turns turbulent its accurate calculati for a long term in the future will coincide with t real development only on the average. T

ocean currents are turbulent and Reynolds number of the atmosphere approaches 10^{10}. Therefore one should be able to determine the turbulence of the Earth's winds and currents correctly.

Climate is just an example of such an averaging, the averaging of not numerical calculations but the averaging of the available information on weather in the past. The notion of climate was introduced by the scientists of ancient Greece. The word itself is also of Greek origin and means "inclination". The Greeks understood that the climate of an area depends on the average inclination of solar rays.

If fact, the main component of the climate of any area, the dependence of averaged weather on the season of the year, is constituted by the conditions of heating this area of the Earth's surface by solar rays, the conditions of illuminance. Consider the variation of the globe's illuminance in a given latitude during a year. The most important thing for the climate is the quantity of energy received by a unit of surface area per day. To find this, however, one should first calculate the alteration of the Earth's illuminance during a day in each season of the year and latitude.

This problem may seem more complicated since one has to obtain more information, but in fact it is easier. The only thing we have to do is to calculate the illuminance of the surface of a rotating sphere by a point source of light. This problem is easily reduced to a purely geometrical one.

The illuminance, i.e. the light power, received by a unit area, as you know, is inversely pro-

portional to the squared distance to the source. Besides, it is proportional to the cosine of angle α between the direction to the source of light and the normal, the perpendicular to the given area:

$$E = s_{\odot}\left(\frac{a_{\oplus}}{r}\right)^2 \cos\alpha,$$

where $s_{\odot} = 1.36 \times 10^3$ W/m² is the solar constant. According to its definition this very constant is the illuminance of a unit area by the Sun's straight rays; the area being at a distance of $a_{\oplus}$ from the Sun. The only difference is that the solar constant is expressed not by luxes, the units of light, but by units of energy. They are convertible into one another with the help of the following relationships:

$$1 \text{ lx} = 0.001471 \text{ W/m}^2; \quad 1 \text{ W/m}^2 = 679.6 \text{ lx}.$$

The distance to the Sun r changes during a year very insignificantly: at eccentricity equal to 0.0167, the distance difference amounts to 3.3%. This means that the total difference in the Earth's illuminance in the intervals between its positions at perihelion and aphelion is almost 7%. This is a significant value. We shall take it into account when calculating total energies received by different latitudes during a day.

However, the main contributor to the change of illuminance is the variation in the angle of inclination of solar rays with time. The angle between direction to the Sun and the vertical of the given place changes significantly during a day and from day to day in the course of a year.

Therefore let us find first the dependence $\alpha(t)$.

It would be convenient to assume that all the angles in the problem change in a wide range. In this case the formula of illuminance is certainly untrue for negative values of the cosine, i.e. when $\pi/2 < \alpha < 3\pi/2$. It is easy to see that this means night at the given place at the given time. In this case the illuminance equals zero.

The latitude of the place will be counted, as usual, from the equator but we shall take negative values of φ for southern latitudes. Thus the geographical latitude would change within the limits $-\pi/2 < \varphi < \pi/2$. Longitude λ, as we know, is counted from the Greenwich meridian. We shall not divide it into "to the west of Greenwich" and "to the east of Greenwich". There is a simpler way. Let us assume that longitude λ changes within the limits between 0 and 2π, between 0° and 360°, the positive direction will be to the east. Then, for example, $\lambda = 2\pi/3$ will mean longitude 120° east and $\lambda = 3\pi/2$ will be longitude 90° west.

The problem of dependence $\alpha(t)$ will be solved by two steps. First, we shall find the yearly variation of the angle γ between the axis of the Earth's rotation and direction to the Sun. In doing so we shall assume that the Earth's orbit is approximately a circle. Consider Fig. 48. It is more convenient to count time t not from the New Year but from the moment of winter solstice. In this case the angle of the Earth's travel round the orbit from that moment equals $2\pi t/T$. Period T here is certainly equal to the tropical year.

The unit vector* **s** of the direction from the Earth to the sun, as indicated in the figure, has components $(-\cos 2\pi t/T,\ -\sin 2\pi t/T,\ 0)$. The unit vector **m** of the direction of the axis of rotation is inclined by angle ε to the z-axis. Its projections onto the axes of coordinates equal

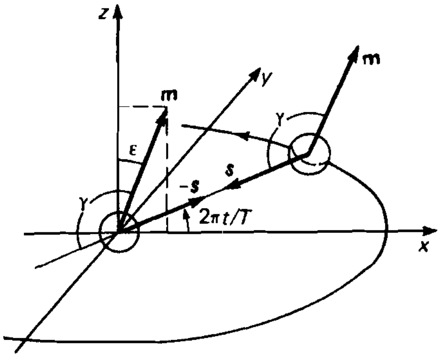


Fig. 48. Demonstration of the formula for the angle between the Earth's axis and direction to the Sun.

$(\sin \varepsilon,\ 0,\ \cos \varepsilon)$. The cosine of the angle between two unit vectors equals their scalar product, the sum of the products of individual projections:

$$\cos \gamma = (\mathbf{sm}) = -\sin \varepsilon \cos 2\pi t/T.$$

Let us check this result in particular cases. This will convince those who are still unable to understand its derivation.

* A unit vector is a vector the square of whose length, i.e. the sum of squares of all projections, equals unity.

1. $t = 0$; $\cos \gamma = -\sin \varepsilon$; $\gamma = \varepsilon + \pi/2 = 113.5°$. In winter the axis of rotation, as indicated in Figs. 12 and 48, does make an obtuse angle with the direction to the Sun.

2. $t = T/4$ or $t = 3T/4$, the vernal or autumnal equinoxes. In this case $\cos \gamma = 0$, the axis of rotation is perpendicular to solar rays.

3. $t = T/2$, the summer solstice; $\cos \gamma = \sin \varepsilon$, angle γ reaches the minimum value of 66.5°.

This simple formula produces the dependence $\gamma(t)$ with an accuracy sufficient for our purpose. It certainly cannot be employed to compile astronomical tables—in that case the ellipticity of the Earth's orbit should be taken into account—but its practical value is doubtless.

Consider now how the inclination of solar rays changes according to the time of the day and geographical coordinates; assume that the dependence upon the season of the year is given by the angle γ. The Earth rotates in relation to the Sun with a period approaching 24 hours. In fact, the solar day equals $P_0 = 86400$ s only on the average because the Earth's velocity at its elliptical orbit is irregular. For example, in January, when the Earth passes the perihelion of its orbit, the solar day is by 30 s longer than P_0. The total deviations of the solar time reach 14 minutes in February and 16 minutes in the reverse direction in October. For simplicity we neglect these deviations and assume that the Earth rotates in relation to the Sun with a constant angular velocity $\omega = 2\pi/P_0$.

Consider Fig. 49. Assumingly, we are interested in the angle of incidence of solar rays in latitude φ in Greenwich meridian $(\lambda = 0)$ at the

moment of the day t by Greenwich mean time This means that the Earth has turned to th angle ωt in relation to the position which occupied at midnight by Greenwich mean tim

The unit vector of the direction to the Sun, makes an angle γ with the z-axis; we know alread

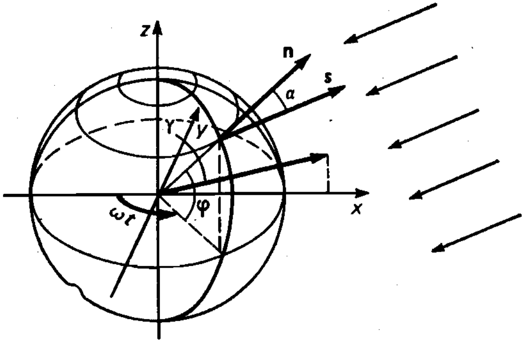


Fig. 49. Demonstration of the formula for the angle o illumination of an area of the Earth's surface by th Sun in latitude φ at the time of the day t.

its value at any season of the year. Therefor projections of this vector onto the axes of coord nates, selected as indicated in Fig. 49, equ ($\sin\gamma$, 0, $\cos\gamma$).

The unit vector of the normal to the point the Earth's surface with geographical coord

* The local time of the day is usually counted off fro midnight. Midnight is 12 hours after the moment of th highest position of the Sun above the horizon, aft midday. The intervals between sunset and midnigh and midnight and sunrise are practically equal.

nates φ and $\lambda = 0$ at the moment of the day t, as indicated in Fig. 49, has components

$$\mathbf{n} = (-\cos\varphi\cos\omega t,\ -\cos\varphi\sin\omega t,\ \sin\varphi).$$

Make up the scalar product of vectors $\mathbf{s}$ and $\mathbf{n}$. Its value equals the cosine of the angle of incidence of solar rays we are interested in:

$$\cos\alpha = (\mathbf{ns}) = \cos\gamma\sin\varphi - \sin\gamma\cos\varphi\cos\omega t.$$

Check now this formula for various specific cases.

1. At midnight $t = 0$, therefore $\cos\alpha = -\sin(\gamma - \varphi)$. The cosine turns out to be negative almost for all latitudes. This means that at midnight the Earth's surface is, as a rule, not illuminated. However, in some latitudes $\cos\alpha > 0$ is possible at midnight as well. This is the area of the polar day. The boundary of the polar day is determined by the condition of the tangential incidence of solar rays at midnight, i.e. $\alpha = 90°$. Latitude φ_1 from which the polar day starts is determined by the equation $\sin(\varphi_1 - \gamma) = 0$. This equation has two solutions in the accessible area of angle variations:

$$\varphi_1 = \gamma \text{ and } \varphi_1 = -\pi + \gamma.$$

It is easy to see that these conditions really produce true boundary latitudes of the polar day: the first is valid for the northern hemisphere in summer, the second takes care of the southern hemisphere between the autumnal and vernal equinoxes.

2. Consider now the inclination of solar rays at midday when $\omega t = \pi$. In this case $\cos\alpha = \sin(\gamma + \varphi)$. The solution of this equation ($\alpha = \gamma + \varphi - \pi/2$) looks general enough, we

17-01137

are not troubled by negative values of α. But cos α turns negative, it means night. A night midday is the polar night. Its boundary latitu φ_2 is also easily derived from the equati $\sin(\gamma + \varphi_2) = 0$:

$$\varphi_2 = -\gamma \text{ and } \varphi_2 = \pi - \gamma.$$

The first expression is true for the southern hem sphere while the second holds for the norther hemisphere in winter.

3. Is there a way to calculate the duration o day and night in any latitude, at any time of th year? The moments of sunrise and sunset are con ditioned by $\cos\alpha = 0$. The duration of night the double sunrise time, consequently

$$\Delta t_{\text{night}} = \frac{P_0}{\pi} \arccos(\tan\varphi \cot\gamma).$$

If the result does not exceed the duration of th day P_0, the remainder $P_0 - \Delta t_{\text{night}}$ is the dur tion of the day. In the opposite case it is pol night. If the number under the argument of th arc cosine is more than one, you have evidentl got into the domains of the polar day.

4. Consider what happens during equinoxes whe $\gamma = \pi/2$. In that case $\cos\alpha = -\cos\varphi\cos\omega t$. Th reveals that the sunrise occurs at 6 o'clock the morning and the duration of the day amounts 12 hours in all latitudes. Day equals night. Th sunset takes place at 18 o'clock*.

* Comparing calculations by this formula to the da of calendars one should bear in mind not only the abo mentioned inaccuracies of the formula, but also th the sunrise and the sunset are indicated in calend not by the center of the solar disk but by its upper ed This increases the duration of the day more than 2 minutes.

5. Now let us move to the Earth equator and assume that $\varphi = 0$. In that case

$$\cos\alpha = -\sin\gamma\cos\omega t.$$

Look carefully: the duration of the day at the equator is exactly 12 hours independently of the season of the year. Solar rays fall vertically at the equator at midday during equinoxes.

You can apply the formula to other interesting particular cases yourself. These include the moments of solstices, the tropical latitudes, the latitudes of polar circles. We have slightly digressed from climatic problems but in order to get through with this formula it should be mentioned that it holds true not only for the zero longitude. The Greenwich meridian was taken as reference longitude on the Earth for no physical reason. Great Britain was the first country to start scientific measurement of longitude by the solar time and the leading observatory of Great Britain was located in Greenwich. The same formula holds true for any other longitude if the local time is substituted in it. It is counted off from the moment of the astronomical midnight at the given place. It is clear that the local time t_l is the Greenwich mean time plus the time it takes the Earth to turn by the angle equal to the longitude of the place:

$$t_l = t + \frac{\lambda}{\omega} = t + \lambda\,\frac{1\text{ hour}}{15^\circ}.$$

The zone time was introduced for the sake of convenience for people living not far from one another. This is the local time rounded in such a way that its difference from the Greenwich

mean time amounted to an integral number hours. The entire Earth's surface was divid into 24 time zones. Their boundaries pass oceans along meridians with longitudes λ_n $7.5° + n \times 15°$, where n is an integral nu ber. In dry land the boundaries of time zones p along the borders of states, rivers, mount ridges or low populated areas.

For example, Moscow is located in longitu 37°. Hence the local time in Moscow leads t Greenwich mean time by 37°/15° hours, i.e. 2 hours 28 minutes. Moscow is in fact position in the II time zone. However, the Moscow ti differs from the Greenwich mean time not by hours. The time difference amounts to 4 hours summer, between April 1 and October 1, and hours in winter. The similar time shifts are i troduced in many countries of the world for ec nomical reasons.

Return now to the illumination of the Ear by the Sun. We know now the dependence of t angle of incidence of the solar rays on the ti of the year, time of the day, and the latitude the place. Hence the average daily distrib tion of the flux of solar energy over the Earth surface can be calculated. The cooling down night is insignificant due to the great heat c pacity of the Earth's surface, especially of i areas covered with water. Therefore, the sol energy received in a day is the principal cha acteristic of the climate of the given latitude. I change with the season of the year dictates the cl matic progress of seasonal changes.

We have to average the formula for the cosi of the angle of incidence over the local time. Th

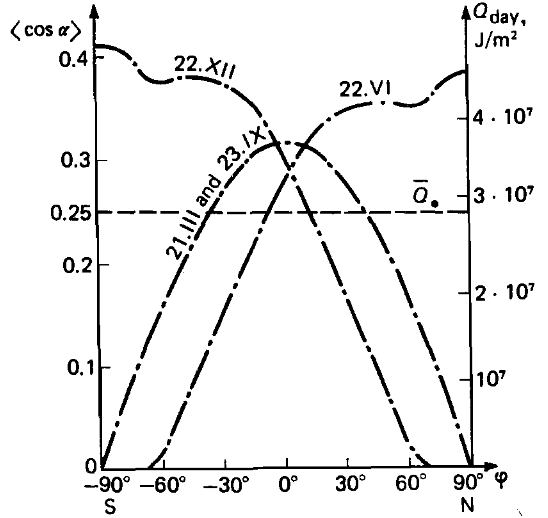


Fig. 50. Latitudinal dependence of one-day light power coming to the Earth during equinoxes and solstices.

averaging is easy for the case of the polar day when second variable term of the formula amounts to zero on the average. Hence the solar energy received by the areas of the polar day in twenty four hours equals

$$Q = P_0 s_\odot \left(\frac{a_\oplus}{r}\right)^2 \cos\gamma \sin\varphi.$$

Calculations for an arbitrary latitudes are more complicated since one should bear in mind that illuminance at night is not negative. These calculations can be performed by those who are aware of integration. The results are given in the form of diagrams in Fig. 50 and 51. In the course

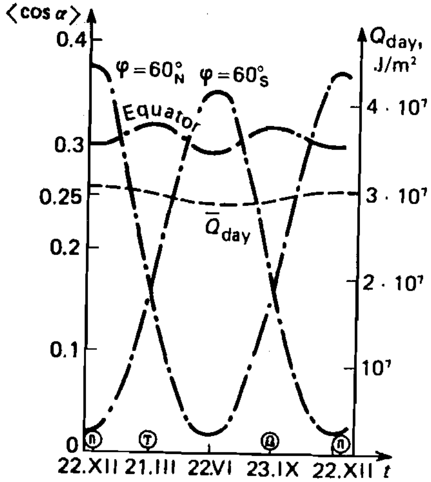


Fig. 51. Yearly dependence of one-day light power com-ing to the equator and latitude 60°.

of plotting the slight change in the distance t the Sun over the year was taken into accoun

Pay attention to the surprising result of th calculations indicating how much energy is re ceived by the polar areas in summer: more tha by the equator, more than by the summer tropic This is not a mistake. This happens because th duration of illumination is twenty four hour during a polar day and the cosine of the angl of incidence is quite significant.

Note also that the June curve in Fig. 50 is no symmetrical to the December one. Curves indi cated in Fig. 51 for northern and southern hemi spheres cannot be produced from one another b

a half-year shift. This happens because in the northern hemisphere we are closer to the Sun in winter than in summer. Recall that the Earth passes its perihelion now at the beginning of January. The areas located in northern latitudes receive on the average the same amount of heat as the similar areas located in southern latitudes but this energy is distributed over seasons somewhat differently. This asymmetry, together with uneven distribution of dry land and oceans over hemispheres, leads, as we shall see below, to interesting and significant peculiarities of the Earth's climate.

Having studied the Earth's radiation equilibrium, we calculated by the average flux of solar energy, the Earth's average temperature, i.e. the temperature of the planet's radiation. At first sight here, too, one could try to estimate the equilibrium temperatures of sections of the Earth's surface if the distribution of the solar energy over latitudes and seasons of the year were known.

This, however, is quite a difficult task. The significant change in the albedo of the Earth's surface with latitude should be taken into account. The albedo, the averaged coefficient of surface reflection, increases near the poles both because there is more snow and ice in those areas and because atmosphere reflects better the oblique solar rays and light comes to the polar areas tangentially. In addition to that it is clear without any calculations that such an estimate would produce a result significantly different from the reality. No solar energy comes to the polar areas during the long night, therefore temperatures there would be extremely low. This, however,

does not happen. Temperature in Arctic and Antarctic regions is only slightly lower than the radiation temperature of the Earth and are far from the absolute zero.

The reason for this is easily comprehensible. The relatively high temperatures of the polar

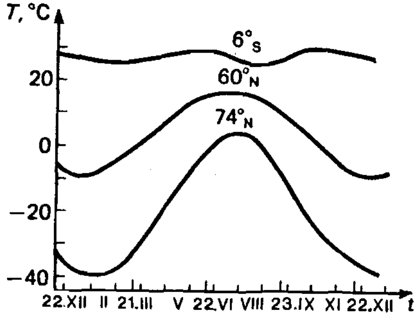


Fig. 52. Yearly dependence of average temperature in different latitudes.

latitudes in winter are supported by the heat exchange continuously occurring on the planet. Thermal energy is continuously transferred from the tropical latitudes to the polar. This transfer is performed by atmospheric winds and ocean currents. In the final analysis, the winds and currents themselves are caused by the temperature difference between high and low latitudes.

Consider now the real changes in average temperatures in some latitudes of the Earth during a year (Fig. 52). You see that these dependences look like the curves indicated in Fig. 51, the distribution of the light energy of the Sun over geographical latitudes. However, the tempera-

tures of the polar areas during the polar night really do not approach the ultimate low values: the thermal energy is redistributed over the Earth's surface.

Note also that the minima and maxima of temperature curves lag behind approximately by one month compared to the extrema of the solar illuminance. This happens due to the thermal inertia of the ocean, dry land, and atmosphere, their heat capacity.

The ocean has the highest heat capacity. We know already that the turbulence and surface currents mix the ocean's upper layer to the depth of approximately $h \sim 100$ m. The specific heat of water equals $c_0 = 4.18 \times 10^3$ J/(kg K). Therefore the total heat capacity of the active upper layer of the ocean covering 70% of the Earth's surface $4\pi R_\oplus^2$ is

$$C_{oc} \simeq 0.7 \times 4\pi R_\oplus^2 h \rho_0 c_0 \simeq 1.7 \times 10^{23} \text{ J/K}.$$

Consider the time Δt during which solar rays illuminating the Earth with power $\pi R_\oplus^2 s_\odot = 1.75 \times 10^{17}$ W would heat the ocean by, say, $\Delta t = 3$ degrees:

$$\Delta t \sim \frac{C_{oc}\Delta T}{\pi R_\oplus^2 s_\odot} \sim 3 \times 10^6 \text{ s} \sim 1 \text{ month}.$$

As follows from Fig. 52, the estimate approaches real values: temperatures lag behind the illuminance approximately by a month and the monthly temperature increase in spring in middle latitudes is about 3 degrees.

The modern term climate implies a notion much more sophisticated than the idea of an-

cient Greeks. The Earth's climate is the entire combination of time-averaged weather data: temperature, pressure, humidity, directions of winds and currents at all points of the planet for each day of the year. There is however a question which time interval should be taken to average weather in order to obtain the climate characteristic.

For definiteness let us treat temperature. Temperatures can be averaged over an interval which is short compared to the seasons of the year and during which conditions of the solar illuminance cannot change significantly. Usually, this period is limited by ten days, one third of a month. Average ten-day temperatures demonstrate a regular development approaching the curves indicated in Fig. 52.

However, a comparison of these average ten-day temperatures in different years would demonstrate inaccurate coincidence, a divergence may reach ten degrees. Their average value for, say, ten years would differ from the average value for another decade only by several degrees. Let us average temperatures over an even longer period: one third of a century. We shall see that the proximity of such average values would further increase while the range of divergence would further reduce.

Let us take average values for 100 years. This brings about a new notion. The difference between adjacent hundred-year values and averages would remain approximately the same: it would not become less than fractions of a degree. This, however, we refer to as a slowly changing climate.

Why are the deviations of average ten-day temperatures from the climatic dependence considered as random phenomena, weather fluctuations, while the variations of hundred-year and thousand-year weather values are regarded as climatic changes? The matter is that changes even in ten-year averages have different signs at different points of the Earth and the sum of these variations over the planet approaches zero. However, the behaviour of hundred-year averages is almost the same over the entire Earth. For example, between the 16th and 17th centuries it was by several degrees colder than now and between the 11th and 12th centuries it was slightly warmer on the entire planet. This is the reason why the alterations of hundred-year average values are attributed to climatic change and the climate itself is considered a global characteristic.

It is interesting to know that the time interval during which the climate can be considered as constant approximately coincides with the duration of human life. Thus, an intuitive approach to weather treats it as something fluctuating about the permanent sinusoid of the seasons of the years, about the constant climate. Only the scientific approach to climate was able to reveal its variability.

The first evidence of long very cold periods in the history of the Earth was provided by the study of glaciers. Snow falling out high in the mountains does not melt even in summer. Being compacted it forms ice. This ice starts to flow down gorges under the action of its own. weight just like a river but millions times slower

The flow velocity of a glacier is several hundred meters per year, i.e. several meters per day. A glacier carries with it stones, the fragments of rocks. A glacier melts upon coming down to a valley but the terminal of the glacier remains approximately constant for years because new ice masses are continuously supplied by mountains.

A glacier deposits heaps of round boulders which it has carried near its edge. These deposits are called end moraines. It is impossible to mistake them for any other geological formations. The position of end moraines allows us to estimate how far had the glacier tongue extended down in the past.

It has turned out that a number of end moraine profiles can be traced around the Alps. These moraines are located considerably lower than the present edges of glaciers. Later end moraines, the age of which approached that of the alpine moraines, were found in Middle-Russian Upland, North Europe, and North America. Thus, 125 thousand years ago a glacier extended from Scandinavia deposited an end moraine near Moscow and 250 thousand years ago a glacier got to the Dnieper downstream of Kiev.

Several dozens of glacial periods in the history of the Earth have been detected. The occurrence was quite irregular: the intervals between glacial periods varied between forty thousand years and several hundred thousand years. Between the glacial periods the climate returned more or less to its present state. The last glacial period deposited its end moraines only twenty thousand years ago.

Yet is there any confidence in the assumption that the glaciation occurred simultaneously all over the planet, are the end moraines the witnesses of the changes in the Earth's climate? Yes, there is. The total mass of ice accumulated in glaciers on dry land was extremely great, its volume twice exceeded that of the present glacial shields of Antarctica and Greenland. All that water was extracted from the World Ocean due to which fact its level dropped many times by more than 100 m below the present one. This is proved by independent geological evidence.

Presently the level of the World Ocean is changing comparatively slow. For the last six thousand years the level alterations did not exceed 3 m. The water level in the ocean rose by merely 10 cm from the beginning of the 20th century. Certainly, these fluctuations are also connected with changes in glacial cover but on the whole a comparison of these fluctuations to enormous level alterations in the past allows us to consider the climate of the last 6000 years, the climate of our civilization, as constant.

But are there any glacial periods to be expected in the future?

3. Causes of Climatic Changes

Climate changes not in individual regions of the Earth but all over the planet. Therefore the cause for its changing should be also global and act upon the entire Earth. This makes clear that the major contribution to the climate formation is made by the thermal equilibrium of the planet. The equality of the fluxes of the solar luminous

energy absorbed by the Earth and infrared thermal radiation released by it makes it possible to calculate the radiation temperature of the Earth:

$$T_{\oplus}=T_{\odot}\sqrt{\frac{R_{\odot}}{2a_{\oplus}}}(1-A)^{1/4}.$$

It is lower than the mean temperature of the Earth's surface but approaches it and determines it together with the thermal equilibrium of the troposphere.

Consider carefully which factors of this formula can change and how fast they can do it. The average distance from the Sun to the Earth, the long semiaxis of the Earth's orbit, is not disturbed by attraction of other planets. The only conceivable change in it can be related exclusively to the change in the Sun's mass. As mentioned above, the Sun's mass decreases in the course of radiation due to mass defect and increases when the Sun is hit by comets. The first effect is insignificant even in the cosmological time scale, the magnitude of the second is difficult to estimate but it also could scarcely cause a notable change in the distances between planets even for billions of years.

Radius $R_{\odot}$ and temperature of the Sun's surface $T_{\odot}$ have entered the formula from luminosity $L_{\odot}$, the luminous power of the Sun. A change in it could be related to the evolution of the composition of the nuclear combustible. The specific time of this process has been estimated. It is extremely long: about 10^{10} years. Therefore, its effect on the Earth's climate is possible only

in time intervals comparable to the age of our planet. Accurate calculations of the Sun's evolution in the process of nuclear burning of hydrogen indicate that 4 billion years ago the Sun's radius was equal to 0.93 of its present radius and the temperature of its surface was by 3 to 4% lower than the present one. This means that the Sun's luminosity was by 25 or even by 30% lower those days than today.

Such a minor difference results, however, in a qualitative paradox. In that case the formula for the radiation temperature produces such a low temperature of the Earth that even with regard to the greenhouse effect its complete glaciation seems inevitable. However, this conclusion contradicts the data on the geological history of our planet. In addition, completely glaciated Earth could never transit to the present warm climate because of the high albedo of ice.

The "small Sun" paradox can be solved if the thermal flux from the interior was taken into account in the thermal balance of this planet in the first billion years of its existence. In the section "Why is the interior of planets hot?" which treated the Earth's formation, we have assessed the average heat flux in the first billion years at 10^{15} W. This additional power supplied by incandescent interior of the Earth through volcanic eruptions is sufficient to maintain positive Celsium temperatures at least in the tropical zone of the Earth.

The surprising constancy of the Earth's climate, keeping temperature in the ocean within the range of merely 30° for billions of years was doubtless vital for life to appear. How low then

is the probability of life on other planets, at oth
stars?! Finally comes the Earth albedo A. Is
changeable? It certainly is. The matter is th
the value of A is determined by the coefficien
of reflection of the atmosphere, ocean, and dr
land. Clouds of the atmosphere, snow cover glacie
of the dry land, and floating ice in seas refle
light most intensively. But they themselves depen
not only upon climate but also upon weather an
season of the year! Thus, the climate can ac
upon itself and its changing depends upon it
own state at the given moment.

The effects of that kind can be of two prin
cipally different types. A pendulum is an exam
ple. A minor deflection of it causes a returnin
force. The pendulum's acceleration is directe
inversely to displacement. Pendulum's equilib
rium is stable. But turn it upside down so tha
the point of suspension were beneath the cente
of masses. This is also an equilibrium but it
unstable: a minor deflection would cause a forc
accelerating the pendulum from the equilibriu
position. The system constituting ocean, atmo
sphere, and dry land is much more complicated tha
a pendulum. But at least one of its instabilities
visible at first sight. Let the climate slightly coo
down. This would yield an increase in glacier
an increment in their area. The snow and ic
surface has a high reflectivity: 0.7 to 0.9. Ther
fore, the Earth's albedo increases on the whol
proportionally to the cooling down. Consequen
ly, the thermal power heating the planet de
creases and it is further cooled down.

The other elements of the system also contr
bute to instability. The level of the ocean dro

during glacial periods, its surface decreases by the area of the shelf, i.e. approximately by 5%. Besides that its ice cover grows. It should be taken into account that the albedo of the uncovered ocean is low, only about 0.1, the ocean easily absorbs solar energy. Thus the ocean's contribution to the Earth's thermal balance is also less during the glaciation.

What about the atmosphere? The lower the temperature, the lower evaporation and saturated air humidity. This brings down the concentration of water vapour in the atmosphere. It is mainly water vapour which keeps heat near the Earth's surface and prevents it from immediate radiation into space. Another minor component of air capable of absorbing infrared radiation is carbonic acid. But the lower the temperature, the better it dissolves in ocean water, which decreases its content in the atmosphere and the atmosphere becomes even more transparent to the thermal radiation.

After all things mentioned above it is unclear why the Earth's glaciation is not permanent and what miracle has been maintaining for six thousand years warm climate. The answer is that there are other processes supporting equilibrium. They ensure the relative stability of the climate. Here is one of them. The Earth's ocean and atmosphere make up a large heat engine. Convection induces winds and currents. Temperature decrease in high latitudes, where the major part of glaciers is concentrated, increases temperature difference between the equator and the pole. An increment in temperature difference between the heater and the cooler of a heat engine increases its efficiency.

18—01137

This enhances the equatorial convection and he transfer by currents from the tropics to the nor and to the south. Currents heat the polar areas a climate returns to a stable state.

Yet the glacial climate is also relatively stabl The duration of each of glacial periods reache dozens thousand years. It is hard to say how th Earth got out of glaciations. It is known that th melting of immense glaciers which had covere entire northern Europe, Asia, and America occurre relatively fast: approximately in a mere thou sand years. The answer is difficult because w know very little about the distribution of atmo spheric winds and ocean currents during the gla ciations. The humans of the Stone Age left us n evidence on that subject and interpreting geo logical data is not an easy task. But apparently i will be soon possible to calculate climate o other epochs with the aid of computers.

One of the possibilities for climate to actuall pass from the glacial state to the warm state an vice versa is a change in the pattern of currents It is known that meandering marine currents ca alter their major direction for ten years. If such a thing happened to a large ocean current, a chang in climate could be rather dramatic. Here i an example of such an event.

Six million years ago there was no Straits o Gibraltar separating the Mediterranean Sea from the Atlantic Ocean. The Mediterranean was con nected with the Atlantic by a more shallow chan nel called the Strait of Riff. A valley remaine from that channel is presently located on th territory of Morocco. Apparently, glacial period occurred also when the ocean's level dropped by

hundred meters. One of such level drops turned the Mediterranean into an isolated basin separated from the World Ocean like the modern Caspian Sea. The Mediterranean Sea is located slightly northward of the desert belt surrounding the Earth. Due to that neighbourhood the evaporation of water is very intensive in this sea and rivers failed to compensate that loss of water. The entire Mediterranean had dried out!

The bottom of its deep part was covered by a layer of sea salt a hundred meters thick. Rivers falling into the Mediterranean—the Nile, Rhone, and Po—had cut canyons one to two kilometers deep and up to thousand kilometers long in soft underlying rocks: both where these rivers flow presently and across the bottom of the former sea. The rivers had fallen into small drying out salt lakes located three kilometers below the ocean level.

When glaciers retreated, the ocean level elevated again, its water started to flow over the Riff valley. The powerful stream of ocean water filled the Mediterranean again in a short time of several hundred years. Rivers gradually filled with sand and silt their canyon-beds and returned to the old river beds. Yet the character of sediments under the present beds of these rivers made it possible to identify the locations of their river beds in the time when Mediterranean had dried up. Such an ancient canyon under the Nile's valley was discovered by the Soviet geologist I. S. Chumakov in the course of construction of the Aswan High Dam. He, too, suggested the hypothesis of the Mediterranean's drying up.

The layer of salt remained at the bottom the newly filled sea. Having been covered sediments this salt did not dissolve in the ne ocean water. Several expeditions of the Americ research ship "Glomar Challenger", which carri special equipment for drilling the sea botto have examined rock samples extracted from bo holes to reveal that the drying up and filling the Mediterranean had repeated many times. A least eleven layers of salt alternated by sedimen make up a stratum two kilometers thick beneat the bottom of the existing Mediterranean Se Apparently, significant drops and elevations the world ocean level took place eleven times an the Mediterranean dried out completely eleve times before the Straits of Gibraltar was expose It happened as follows.

Five million years ago a terrific earthqua cracked the Earth's crust and moved it apart b dozens of kilometers from the Mediterranean t the middle of the Atlantic Ocean. The Atlant waters thundering like a hundred of Niagara Fal rushed into the Mediterranean through that fra ture several kilometers wide and more tha thousand meters deep. The Straits of Gibralt having been exposed, the Mediterranean wa filled by the cold ocean water populated by orga isms characteristic for the ocean depths. Ho ever, two million years ago the Straits of Gibra tar shoaled to its present depth of five hundre meters and water at the bottom of the Medite ranean warmed up to the temperature of 12 ° (the currents therein were treated in the previo chapter). All deep-water organisms of the Medite ranean Sea failed to survive the warming-u

and up to this day the sea remains a biological desert below the depth of 2.5 kilometers from the surface.

In the final analysis, all the salt buried at the bottom of the Mediterranean had been extracted from the World Ocean. Owing to this, the ocean's saltiness has decreased quite significantly. It is easy to calculate that the decrease amounted to $4\,^0/_{00}$ after which the salinity was slowly increasing until it reached the initial level of $35\,^0/_{00}$. The decrease in ocean's salinity reduces the melting of ice covering its polar areas. Therefore, the boundary of sea ice passed, apparently, further south, 6 to 5 million years ago, than the present line. The Earth's albedo was slightly higher than now. It is most likely that the planet's climate should have been colder on the average at that time than both the climate of today and the climate before the Mediterranean's drying out. We would remind you that all those events have taken place very recently compared to the age of the Earth but much earlier than the last glacial periods.

In contrast to the above mentioned, the ocean's salinity increases during the usual glaciations because glaciers accumulate fresh water. The ocean level drop by 100 m increases its salinity by $1\,^0/_{00}$. The change in salinity and alteration of the coast line can bring about a change in its major currents. This, in turn, causes a powerful inverse action on the climate. Why is the system of World Ocean currents so sensitive to a change in its salinity? The answer is not a simple one.

The temperature of ocean water at the depths lower than 1.5 kilometers falls down to 2-3 °C

almost independently of latitude. This resul in the accumulation of an enormous store of col inside the World Ocean. Judge for yourself, th is almost a paradox: the mean temperature o the Earth's surface equals 15 °C, the temperatu of the Earth's interior is also high (at the dept of 7 km it always exceeds 200 °C), and it is mi imum at intermediate depths both in dry lan and in the ocean. By analogy with the greenhous effect maintaining near the Earth's surface temperature rather high compared with the plan et's equilibrium radiation temperature, thi phenomenon, namely, the cold of the ocean depth and the subsurface cold of soil, could be calle the cellar effect.

Why does the heat conduction fail to equaliz the temperature of deep ocean water by heat fluxe from below and from above? The reason for thi is identical to that owing to which the tempera ture in caves is always lower than the mean yearl temperature on the surface. This reason is the au tumnal-winter convection of ground and ocea water.

During the summer solar rays warm up onl the top layer of the surface: about a hundred me ters of ocean and sea water and merely 1 to 2 m of soil. The density of warm surface water bein lower than that of the unheated water underneath the mixing of them in spring and in summer is very slow process. It is quite different in autum and in winter. The cooled (especially at nights) surface water becomes heavier than the underlyin water which has warmed up during the precedin warm season. The convection starts. The sinkin of cold water replaced by elevating warm wate

results in a heat flux directed upwards. It is significantly more powerful than heat fluxes produced by heat conduction.

However, there are important differences between the autumnal-winter convection of ground water and the cooling of ocean depths. Firstly, the density of almost fresh ground water has a maximum at 4 °C, therefore, the convection of ground water and fresh water of lakes ceases as soon as the temperature of the surface falls below +4 °C. This is why the temperature in caves and at the bottom of deep lakes holds at approximately that level in midlatitudes. In tropics the depressed ground water temperature is determined not by the contrast between summer and winter but by the difference between day and night temperatures. The ocean water with salinity of 35 ‰ has no temperature maximum of density. The colder the water, the heavier it is, down to temperature of —2 °C when ice crystals appear in it.

Secondly, the autumnal convection in soil and lakes has small horizontal dimensions. The winter sinking of cold ocean water (downwelling) by no means results in water elevation at the same place but produces large-scale thermohaline currents and upwelling at very long distances. Consider again Fig. 40.

This convection has resulted in the accumulation of enormous negative thermal power in ocean depths, negative because the average ocean temperature, $T_{oc} = 3.5$ °C, is lower than the average temperature of the Earth's surface, $T_{sur} = 15$ °C. You see that nature has invented a way of "sweeping under the carpet" the polar

cold hiding it in ocean depths. Let us estimat this store of cold:

$$Q \simeq m_{oc}c_0 (T_{sur} - T_{oc}) \simeq 6 \times 10^{25} \text{ J}.$$

It is clear that even a minor disturbance of th established system of currents can mix a part o deep water with surface water and lead to a sig nificant cooling of the climate. It is interesting t compare the store of cold in the World Ocea with the energy spent on the melting of glacie of the last glacial period. At that time the ocea level was lower than the present one by approxi mately $h = 80$ m. Consequently, the mass of ice covering the dry land equalled the mass of wate in the upper ocean layer of the same thickness i.e. $m_1 \simeq 0.7 \times 4\pi R_{\oplus} h \rho_0 \simeq 3 \times 10^{19}$ kg. Let u neglect the glaciation of the ocean itself. To melt that mass of ice, an energy of about 10^{25} J wa required. Such a thermal energy is received b the Earth from the Sun in 2.5 years but it is spent as you already know, on the infrared radiatio of this planet. However, this energy is less bu of the same order of magnitude that the presen thermal energy deficit of the ocean, its store o cold. Therefore, any intensification of the dee ocean water mixing with the surface water re sults in the cooling of the climate and a ver intensive mixing of them can bring about even glacial period.

Besides the permanent wind currents an al most regular contribution to the mixing of ocean water is made by hurricanes (typhoons). These moving deep cyclones (the lowest recorded pres- sure near the surface was 0.859 of the normal) occur 6 times a year on the average (from 0 to

13 *times*) in western tropical areas of the Pacific and Atlantic oceans. A hurricane wind with 50 m/s velocity drives 30-meter-high waves; the kinetic energy of a hurricane reaches 10^{19} J. The liberation of this energy and its conversion into heat, however, does not warm up the ocean surface but cools it. Hurricanes mix the tropical ocean intensively down to the depth of 200 to 400 m in an area about a hundred kilometers wide. The water in a track of such hurricane remains colder by several degrees long after the hurricane is over.

Volcanic eruptions produce a short-term effect on the climate. Dust and smoke which they eject to a high altitude are carried by winds of the upper troposphere around the entire Earth. In some cases a smoke column from an eruption reaches even the stratosphere. A total precipitation of this dust takes years. The dust disperses the solar light due to which the Earth's albedo temporarily increases.

Thus, on August 27, 1883 the eruption of Krakatoa took place to become the most powerful volcanic explosion in the history of humanity registered on the Earth's surface. A tsunami wave caused by the fall of a broken-away part of Krakatoa island travelled around the globe. Unusual optical phenomena were observed in Europe already in the end of November 1883: the sky remained purple for several hours during sunsets. That was an effect of the dispersion of the solar light by a layer of dust injected by the volcano into the stratosphere. For a number of successive years the weather on the entire Earth was colder than usually. Note that in the given case the cool-

ing could have been caused not only by the temporary increase in the Earth's albedo, but also by a partial mixing of ocean water by tsunami.

Once a hypothesis was discussed which stated that spontaneous climatic coolings in the past could have been the result of water mixing caused by falls of very large meteorites into the ocean. This reasoning, however, cannot be applied to the explanation of multiple glacial periods since the probability of a fall of a meteorite, which could disturb the ocean more than the Krakatoa explosion had done, is extremely low.

Note one more consequence of the cellar effect. When we discussed in Chapter IV the Earth's thermal balance, we pointed out that we have failed so far to make sure that the heat flux of the infrared radiation going away from the Earth exactly compensates the thermal power brought to the Earth by solar rays. In principle, a situation is possible in which these fluxes would become equal on the average but only within the time span of about several tens of thousand years. In other words, the possibility cannot be excluded that during warm interglacial periods the warmed-up atmosphere irradiates energy into space by a fraction of per cent more than the Earth receives from the Sun, and the ocean depths accumulate gradually a store of cold. In contrast to that, the cooled atmosphere releases into space during glacial periods a flux of energy less than the present one and the changed system of currents can accumulate the lacking energy in the shoaled but still deep ocean.

According to the modern idea both states of the Earth's climate, the warm and the glacial,

approach a stable condition. A transition between them is retarded and hampered by the thermal inertia of formation and melting of glaciers and considerable mechanical energy required for a change in ocean currents. Such transitions can be due to many causes. Moreover, it is not necessary that the external action be a powerful one and notably alter the flux of thermal energy coming to the Earth. The duration of the external action is much more important. In this connection the effect of the Sun's variable nonthermal radiation upon the climate cannot be excluded. Eleven years of the solar cycle is too short a period for a disturbance of the ultraviolet zone of the solar spectrum to produce any effect on the Earth's climate. But, for example, many scientists think that 70 years of the Maunder minimum can well account for the minor cooling of the climate in the 17th century.

There is one external effect which may seem weak at first sight, but which continues for a long time—tens and hundreds of thousands of years. It is probably not this effect which changes the climate from warm to glacial condition and back, but rather it determines the time scale of these changes. It is not a spring but rather a pendulum regulating the operation of the Earth's climatic clock.

The reason is the change in the eccentricity of the Earth's orbit round the Sun and the precession of the axis of the Earth rotation jointly influencing the Earth's climate. An astronomical theory of climatic alterations was originated by the outstanding Yugoslavian scientist M. Milankovitch in the 20s of this century. The theory pro-

vided a possibility to calculate the times of pas glacial periods. The ages of a dozen of forme glaciations coincide with dates offered by tha theory to an accuracy of the ambiguity of geo logical dating. The same theory allows us to an swer the question when the next glaciation o the Earth is to be expected.

4. The Cold "Breath" of the Ecliptic

Milankovitch's reasoning was as follows. Th eccentricity of the Earth's orbit changes unde the action of minor disturbances of other planet (Fig. 11). It may reach the value of $e_{max} =$ 0.0658 which is not a low one. The characteristi period of eccentricity alteration is about 100 thou sand years. In addition, the period of precessio of the axis of rotation is 26 thousand years an the angle of inclination of the axis of rotation t the plane of ecliptic also varies with a perio of 41 thousand years (Fig. 21). Therefore, th conditions of the illuminance of our planet b the Sun change significantly during such time which approach by the order of magnitude th times of change of glacial epochs. Is it possibl that the combination of these astronomical phe nomena had been causing the multiple glaciation of the Earth?

Consider first how the illuminance of the Earth by the Sun changes within a year due to the el lipticity of the Earth's orbit. At the perihelio the distance to the Sun equals $a_\oplus (1 - e)$, whic is less than the average distance $a_\oplus$. The powe received by the Earth from the Sun is inversel proportional to the square of the distance. There

fore, near the perihelion it relates to the mean power as $(1 - e)^{-2} \simeq 1 + 2e$. At the moment of passing the perihelion the Earth's surface is illuminated better and it is relatively warmer on the Earth. But the Earth's orbital velocity near the perihelion is higher than the average; it is proportional to

$$\sqrt{(1+e)/(1-e)} \simeq 1+e.$$

Therefore, the planet passes rather fast the section of its orbit, where the illuminance of its surface is higher.

In contrast to that the distance to the Sun at aphelion is long, $a_\oplus (1 + e)$, and the illuminance of the Earth by solar rays is relatively low: it is proportional to $(1 + e)^{-2} \simeq 1 - 2e$. The Earth spends a relatively long time in the section of its orbit distant from the Sun. The velocity of its motion there is not high: it is less than the average and is proportional to

$$\sqrt{(1 - e)/(1 + e)} \simeq 1 - e.$$

The question is: if the mean distance to the Sun is constant and only the orbit's eccentricity changes, does the total amount of heat received from the Sun in a year depend on it? To answer this one should sum the power of solar radiation coming to the Earth in a year. The matter is that there is a weak dependence upon the eccentricity: the annual energy is proportional to $(1 - e^2)^{-1/2} \simeq 1 + 0.5e^2$.

The result is that the greater the eccentricity, the higher the average illuminance of the Earth, but the difference is very small. Thus, we re-

ceive presently, when the eccentricity equal 0.0167, merely by 0.014% more heat compare to the circular orbit. Even provided the maximu eccentricity, the average energy received by th Earth in a year would increase only by 0.2%.

Since we are interested in the average temper ature of the Earth, the extraction of the fourt root from the yearly sum of heat would reduc these minor corrections by a factor of 4. The in fluence of the orbit's ellipticity will be negligibl Thus, we may assume that the Earth receiv from the Sun the same amount of energy eac year independently of the eccentricity of it orbit.

However, note, Milankovitch said, that th solar power varies during a year not so littl the difference between the illuminance of th Earth at perihelion and aphelion relates to th average illuminance as $4e$. Thus, at great eccen tricities it may reach 0.26 which is a drop by quarter of the average value! Therefore, we hav to consider the alteration of total heat receive by the Earth in separate seasons. The answer t this question depends on whether the Eart passes its perihelion in summer or in winter.

You remember that the change of seasons i determined by the angle of inclination of th axis of rotation to the ecliptic, the plane of th Earth's orbit. The angle of this inclination, th slope, ε equals now 23°26′30″. It also slowl changes, varying between the minimum 22° an the maximum 24.5° with a period of 41 thousan years. The latitudes of tropics and polar circle vary by 2.5°. These variations are important fo accurate calculations but they are not very sig

nificant for further qualitative reasoning. It is more important that the axis of rotation itself rotates almost without changing its slope and precesses around a perpendicular to the ecliptic. We have considered this phenomenon when we were treating the design of our calendar. The period of precession equals 26 thousand years. This is the time after which the North Star will again take its northern position, i.e. when the axis of rotation will be directed at it. However, we are interested now not in the period of precession in relation to stars but in the mutual locations of the orbit's perihelion and points of summer and winter solstices.

The perihelion of the Earth's orbit also transits in relation to stars. It moves in the same direction as the Earth itself moves by the orbit but the velocity of its motion is 100 thousand times lower. This means that the perihelion makes a full turn round the orbit in $T_p = 100$ thousand years. The point of vernal equinox transits by the orbit together with the points of summer and winter solstices in the reverse direction with a period of $T_e = 26$ thousand years.

Thus, to find the period T_0 with which the mutual position of the perihelion and the point of winter solstice is repeated, one should add the angular velocities of these points, i.e. add the values inverse to their periods:

$$\frac{1}{T_0} = \frac{1}{T_e} + \frac{1}{T_p}; \quad T_0 = 21 \text{ thousand years.}$$

Let us find when the Earth's perihelion actually coincided with winter solstice. These days the Earth passes its perihelion on January 4 and win-

ter solstice occurs on December 22. The diffe ence amounts to 13 days of orbital motion or t 12.65°. In 21 thousand years this angle will in crease to 360°. Therefore, relatively recently, i 1250, A.D., the perihelion coincided with th point of winter solstice. Next time they will coin cide in 20 thousand years*, the previous conjunc tion occurred 22 thousand years ago and befor that, it took place 45.5 and 69.2 thousand year B.C. These moments are repeated not with strict periodicity since calculations of them ta into account the disturbances from all planet

The point of summer solstice coincided wi the perihelion in the intervals between the events. It took place 11.2, 33.2, 60.1, and forth, years ago.

Imagine now one of the moments of the coi cidence of the perihelion and winter solstic This position is presently being realized to accuracy of ten degrees. *However*, for the sa of expressiveness, let us assume an eccentrici of 0.06. It was such 230 thousand years ago a it will be such again in 620 thousand years. T vernal equinox keeps to March 21, which is t date recorded in our calendar. But the summ solstice will take place only on June 28 and t autumnal equinox will be postponed and f on October 5. The Earth moves slowly along t distant part of its orbit. The summer of the nort ern hemisphere is long, but the Earth receiv little heat and glaciers melt slowly, languidl Then winter comes. It is short, winter solsti takes place on December 28, only 84 days af

* But the orbit will be almost circular then (see Fig. 1

the autumnal equinox. The Sun is close to the Earth, it throws a dazzling light on snow deserts. The snow melts at daytime but at night everything freezes anew.

Exactly such conditions are favourable for the growth of glaciers, exactly they promote the transition of the climate to a glacial state. Milankovitch wrote: "The advent of glaciers is promoted not by a severe winter but by a cool summer."

Let us move over to the southern hemisphere now. In that part of the globe winter is long and severe, and summer is short and very hot. But *the ocean covering the major part of the southern* hemisphere warms up during summer and moderates the forced continentality of climate.

Merely 11 thousand years pass after the described events and the situation becomes quite opposite. The perihelion coincides then with the moment of summer solstice and the eccentricity still remains as great as before . It is the turn of the northern hemisphere to accommodate long and severe winters. Imagine a hard-ringing frost of —50 °C but failing to produce much snow. Then comes summer, short but very hot. Glaciers melt fast and the ocean level elevates. The glacial period is succeeded by a moderate climate.

Conditions in the southern hemisphere are now favourable for glaciation. But there is no room there for glaciers to accumulate with an exception of the Antarctic, bad enough as it is with its permanent ice cover, and a narrow strip of South America. As for the ocean, it easily liberates from ice during long, although cool,

summer. On the whole, the area of glaciers on t
Earth decreases, but average-annual temperatu
increases.

The perihelion of the Earth's orbit is prese
ly close to the point of winter solstice. It follo
from the above reasoning that such a positi
promotes glaciation. If this is true, why is our cl
mate not glacial now? Furthermore, the glaci
tion has been decreasing during the last hundr
years and the ocean level has been elevating.

A qualitative answer is that the eccentrici
is small and continues to diminish. It equa
presently 0.0167 and in a thousand years
will reach 0.0163. Such a value is insufficient
bring the climate out of the relatively stab
warm state. But 20 thousand years ago the ecce
tricity equalled 0.020 and a glacier tongue fro
Scandinavian mountains reached the prese
city of Kalinin at that time. There was also
Baltic Sea then.

The above discussion is purely qualitativ
Milankovitch has managed to translate it in
the language of numbers. He introduced the fo
lowing quantitative technique to compare th
modern climatic conditions to the climate
other times. Now in the northern hemisphere th
average latitude at which the lower boundary
glaciers descends to the sea level, the snow li
latitude, is approximately 65°. Dry land to th
north of that latitude is covered by glaciers eve
in summer but there is not much land within th
Arctic circle. Milankovitch calculated the sol
energy coming to the Earth's surface in latitu
65° North during the summer period of the yea
It is the illuminance of the Earth which caus

advent or retreat of glaciers, the illuminance when it is summer in the northern hemisphere, and 65° is the latitude from which starts the growth of glaciers changing the Earth's albedo. Having calculated the summer quantities of heat which the Earth received in the past in latitude 65° North, Milankovitch found the latitude in which the quantity of heat received presently by the planet equals the calculated results.

If in the past this equivalent latitude were, say, 69°, it means that at that time climatic conditions in latitude 65° North were close to the conditions existing presently in latitude 69°. Consequently, the boundary of glaciation would shift in such conditions approximately by 4° to the south and glaciers would cover vast territories of dry land in Canada, Scandinavia, East Siberia. In such conditions the Earth's albedo would so increase that climate could transit to a glacial state. In contrast to that if at any moment the equivalent latitude was less than 65°, this means that at that time the glaciation of the northern hemisphere and thus of the Earth on the whole would decrease.

Consider now how the equivalent latitudes were changing in the past and how they will change in the future (Fig. 53*a*). The black sections indicate the time periods when the equivalent latitude approaches 69°, it is these periods which are to be compared with geological datings of previous glaciations. These curves were plotted by the Soviet astronomers Sh.G. Sharaf and N.A. Budnikova according to Milankovitch's theory but they employed more accurate information on the chang-

ing parameters of the Earth's orbit and the slo of the Earth's axis.

The equivalent latitude diagram can be co pared to the temperature data placed under (Fig. 53*b*), *indicating* the past temperatures the Caribbean Sea (it is located in tropics). The

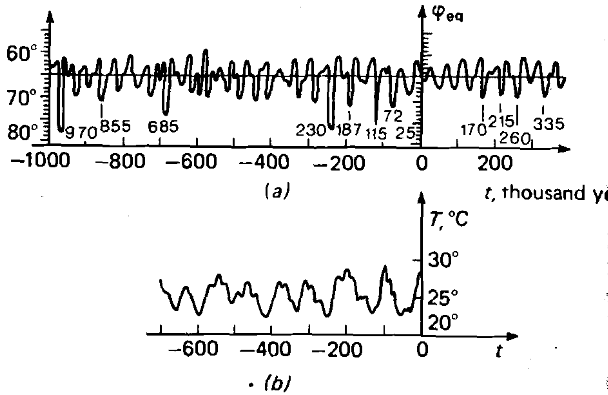


Fig. 53. Alteration of Milankovitch's equivalent la tude (*a*) and temperature of the Caribbean Sea in t last 700 thousand years (*b*).

data were obtained with the help of the radi isotope technique taking advantage of the fa that the content of the heavy isotope ^{18}O in w ter depends upon temperature. It can be seen th even in a water basin located far from glacie in a basin with no radical change of seasons, te perature dropped by several degrees during gl cial periods. This proves that glaciers affe the entire Earth's climate.

Considering these diagrams one should bear in mind that not only planets affect the climate and that it depends not only on the summer illuminance of the 65° latitude. The actual position of glaciers' boundary is significantly influenced by their own past development: the thermal inertia of glaciers is very strong. Thus, after a great glaciation which took place, for example, 115 thousand years ago, the glaciers failed to melt completely even at higher summer illuminance. Therefore, the next glacial periods with lesser amplitudes (the minima of Fig. 53 which took place 72 and 25 thousand years ago) were not so extensive but resulted in significant glaciations. Milankovitch's theory does not forecast climate accurately, it gives the climatic rhythm.

Certainly, the real boundaries of glaciers did not pass exactly along parallels of latitude. For example, there were no extensive glaciers in West Siberia even in rather high latitudes. The main reason for that anomaly is, apparently, the absence of significant mountain ridges in West Siberia. It is the mountains which become the nuclei, the embryos of large glaciers sliding down to valleys. Another cause for the absence of glaciation is the continentality of West Siberian climate, the low precipitation. In such conditions the soil having no snow cover freezes during severe winters down to considerable depth. Thus the permafrost forms. Despite the heat flux from the Earth's interior (treated in Chapter 3) the permafrost may reach the depth of one kilometer. It does not disappear completely even in warm interglacial periods.

5. The Climate and Civilization

The last glacial period terminated about 10 tho sand years ago. This moment corresponds to t last maximum indicated in Fig. 53*a*. The bou dary of glaciers moved further northwards a higher in mountains; the ocean level started elevate rather fast and reached the present o approximately six thousand years ago. Peopl which had lived in the near-tropical zone start to move over northwards and populated vast te ritories of Europe. The most intensive migratio was from Hindustan and Iranian Highland. Th languages of the majority of European peopl have common roots with the language of ancie India, the Sanskrit. This is why one of the maj human races is called Indo-European.

The migration of peoples facilitated contact between them and promoted the propagation o knowledge. Favourable climatic conditions stim ulated cattle-breeding and agriculture. Peopl liberated from the constant struggle against se vere nature got free hands for a spiritual develop ment. First written languages, the main featur of our civilization, appeared approximately a that time. This facilitated the communication o accumulated knowledge not only in oral form from generation to generation, but also in writte form to distant posterity and over long distances The volume of accumulated knowledge start ed to enhance at a pace impossible till then.

Of all things, what do the modern people nee the written language for? Apparently, for exchange of technologies, the techniques of material wealth production. This possibility, how-

ever, was not the incentive for the appearance of a written language in the ancient world. The secrets of trades were safely passed to inheritors within a family up to the epoch of Renaissance between the 15th and 16th centuries A. D. Historically, it was apparently the possibility of recording astronomical and climatic events, of comparing and analyzing them over time periods much longer than the life of one generation, that mainly contributed to the realization of the need for a written language. The effect was the possibility to predict similar events by analogy which is impossible without a written language.

The periodicity of lunar phases certainly can be detected simply by day counting. But even the creation of a primitive 365-day solar calendar in ancient Egypt required many-year observations of the night sky and recording of the Nile's floods, correlating them to the first spring appearance of Sirius, the brightest star. The scope of the mental work done by the ancients should not be underestimated. For example, the forecast of lunar eclipses was successfully performed in the 2nd millenium B.C. by ancient Chinese and Babylonian astronomers. And what about you, the graduates of high school who had studied astronomy? Could you calculate the night of the nearest lunar eclipse with the help of the records of Moon's positions among constellations? I doubt it.

The climatic effect on the progress of humanity was very significant also before the birth of civilization. Here, the major benefactors were not only warm interglacial periods, golden ages of flourishing, but also the cooling which accom-

panied the glaciers' advent. It is quite possib that the variability of the Earth's climate, th alternation of warm and cold epochs with peri dicity of 40 to 200 thousand years, was a nece sary condition for man's maturing as biologic species. Working, thinking, and speaking we stimulated by the need to overcome a crisis, th cold consolidated primitive human collective Maintaining the fire, and later starting fires b friction were invented to fight the cold.

The human migration from Asia and Afric to all continents was facilitated also by the gl cial periods. The depression of ocean level du ing glaciations connected islands with cont nents and continents with one another. Thu for example, disappeared the Bering Strait separating Chukotski Peninsula from Alaska. I is believed that peoples of North-East Asia move to the American continent over an isthmu formed exactly during the last glacial period. Th exodus took place before the Alaska's total gla ciation and emigrants settled then all over Nort and South America. This concept is based on th fact that American Indians are ethnically clos to Mongols and Tatars. At the same time ther is no evidence of human presence on the Amer ican continent earlier than twenty thousan years ago.

Australia was also populated by emigrant from Asia. That, however, took place even ear lier: 50 to 100 thousand years ago. The presentl living Australian natives are ethnically rathe far from the majority of peoples of South-Eas Asia. In the last glacial period the distance be tween Asia and Australia was too great for migra

tion. There was no permanent connection between these continents even during the greatest glaciations which depressed the ocean level by 120 to 150 m. However, such a shoaling was sufficient for presently existing Indonesian islands to get linked with Indo-China while they were separated from Australia and New Guinea by a narrow strait about 100 km wide. We have to assume that even in such ancient times humans could cross water barriers that wide.

Let us, however, return to the beginning of our civilization which was born 6 to 7 thousand years ago. The written language is not its sole distinction from the primitive society. Approximately at the same time the humanity waved good-bye to the Stone Age after having learned to process natural metals (copper, silver, gold). Then people learned to manufacture bronze, build river and sea vessels, and finally the wheel was invented. All the successive progress of humanity follows one and the same pattern: social consequences result from technological developments.

Nevertheless, it was the invention of written language which drew a line between the civilization and the previous existence of humanity. The written language has not only facilitated research into the future with the help of an analysis of the past. Its social significance is that it has reshaped the administration of human collectives, introduced into it a rigid feedback. A written order from above inevitably requires a written report from below. Inobedience to the order causes a punishment on the basis of written laws.

The strengthening of administration had co solidated minor communities consisting of se eral families into large states. The economic ad vantage of unification has given impetus to rapid growth of the Earth's population. Th permanently accelerating growth continues d spite wars and epidemics: the time of population doubling is decreasing monotonically. Thu the warming-up of the climate after the end the glacial period has stimulated the huma progress. But humanity also influences the cl mate.

First signs of climatic changes resulting fro human activities could be detected in some are already at the dawn of civilization. We kno from poems by Homer that mountains of ancie Greece were covered with thick forests. Toda they are bare rocks. Their grass cover was trample down by flocks of domestic goats: goats destro grass more than any other domestic animals.

Another and more significant example is th Sahara. Drilling of the Nile valley has reveale that this desert, the greatest desert on the Eart formed only about 6 thousand years ago. In earli sedimentary layers there is no eolian sand, i. sand deposited by wind, while all upper laye contain unrounded by water sand particles Sahara origin. This means that there was no Sa hara during warm intervals between ancie glacial periods. It is most likely that the dese is the result of man's pasturing cattle on th fragile grass cover.

The albedo of sand deserts is higher than th albedo of areas with vegetation. At the sam time, the air aridity in deserts helps their cool

ing by radiation. Therefore deserts, including the Sahara which occupies 6% of dry land, additionally cool the Earth. These days man plants forests and irrigates arid lands which positively affects the climate.

The Earth's climate has notably warmed up for the last hundred years. The warming-up was most intensive between 1890 and 1940. In the northern hemisphere the average-annual temperature elevated by 0.6 degrees, glaciers were retreating to mountains, the permafrost boundary shifted northwards, the line of permanent sea ice also moved to the north by several degrees. Information on the southern hemisphere is inadequate for a confident estimation of the warming-up. On the whole the Earth's climate has changed in a positive direction since the ocean level has elevated in those fifty years by 10 cm.

Unfortunately, the actual cause of that warming-up is not clear. Its explanation by the results of human activity does not seem convincing because the last 30 years have not brought about a further warming. The average temperature of the Earth's surface has stabilized and in the northern hemisphere it has even dropped by 0.2 degrees. From 1950 the ocean level varied within the limits of ± 3 cm but remained almost constant on the average.

Note that in conformity with Milankovitch's theory the long-term climate tendency is negative now, the alteration of astronomical parameters is favourable for cooling. As follows from Fig. 53*a*, the summer illuminance of the northern hemisphere is presently decreasing and will reach its minimum in 20 thousand years. However, the

nearest minimum is not extensive. It is connec ed with the eccentricity of the Earth's orb which will be almost zero at that time and w know that only a great eccentricity results i glacial periods. Thus, the nearest climatic f ture of the Earth is not pessimistic. "The pos tioning of heavenly bodies", an astrologer woul say, favours a cooling. However, in the last hu dred years the human activity brought about new factor which makes the climate warmer.

The great amount of coal, oil, and gas burn by the humanity started to make a minor bu measurable contribution to the content of atm spheric carbonic acid. In addition to that CO is formed in significant quantities in the cour of cement manufacturing. Carbon dioxide, lik water vapours, absorbs infrared radiation, i. it facilitates the greenhouse effect. Therefor there is a widespread opinion that an increas in the extraction of minerals can bring about notable warming-up of the atmosphere rathe soon.

The fact is that every year we burn with fu about 5×10^{12} kg of carbon extracted from ge logical deposits. This forms 2×10^{13} kg of ca bon dioxide which amounts to 1% of its tot mass in our atmosphere. This, however, does no mean that the content of carbonic acid in th atmosphere increases at the same rate.

The present share of carbon dioxide in the at mosphere equals 3.37×10^{-4} of its total mas Twenty years ago its content was notably les 3.18×10^{-4}. Consequently, the quantity of CO increases annually by 10^{-6} of the total mass o the atmosphere, i.e. by 5×10^{12} kg. You se

that the actually measurable increment in concentration has turned out to be 3 or 4 times less than it could be expected from the calculations of burnt fuel. Where does the rest of carbonic acid disappear then?

The cause of this discrepancy is that carbon dioxide easily dissolves in ocean water. The dissolved carbonic acid is consumed by phytoplankton (unicellular chlorophyll organisms living in the upper layer of the entire ocean) in the course of photosynthesis. Besides that carbonic acid not only dissolves in the ocean but also undergoes a chemical reaction: $CaCO_3 + CO_2 + H_2O \rightarrow Ca(HCO_3)_2$. Chalk, $CaCO_3$, is deposited in enormous quantities on the ocean bottom in the form of dead shells and coral skeletons. This chemical reaction dissolves chalk deposits since calcium hydrocarbonate, $Ca(HCO_3)_2$, dissolves in ocean water much better than chalk. Because of this reaction it is quite difficult to answer the question how much carbonic acid is presently dissolved in the ocean. But we can say with confidence that the ocean is in principle capable of absorbing a quantity of carbon dioxide by far exceeding its atmospheric mass. However, this process is not fast enough. It is only due to the low rate of solution of carbon dioxide and low rate of this chemical reaction, that we record a certain increment in its content in the atmosphere.

Assume for a moment that the concentration of carbonic acid in the atmosphere has doubled. This would result in an elevation of the average temperature of the Earth's surface by 2 to 3 degrees. The warming-up would be especially notable in polar areas, where the average tem-

perature would elevate even by 6 to 8 degrees, si
the entire Arctic Ocean would be free of ice
summer. Alas, the doubling of CO_2 content
the atmosphere will be apparently ne
reached. Firstly, the rate of carbon dioxide abso
tion by the ocean increases with its concentra
in the atmosphere. Secondly, the photosynth
will be more active in a warm climate and p
tosynthesis transforms carbonic acid into orga
life of vegetation both on dry land and in oce

However, we have to care not only for the wa
ing-up of the climate but also for escaping
overheating of the Earth, especially in trop
areas in latitudes of the desert zone. Besides t
a significant warming-up would decrease glac
of the Antarctic and Greenland, the ocean le
would elevate. The resultant flooding would
off and bury under water fertile coastal areas
dry land. On the whole the economic consequen
of a warming-up would be rather negative
the Earth.

It is interesting to estimate how soon wou
the direct heat release by the global power en
neering produce a notable effect on the clima
In this case one should certainly calculate on
the energy produced by combustion of mine
fuels and the atomic energy. Firewood combu
tion and burning of agricultural wastes, dire
conversion of solar energy into other forms, wa
power stations and wind motors use directly
indirectly the energy of the solar radiation. T
solar radiation has been already taken into a
count when we treated the Earth's thermal equ
librium. Here we deal only with an addition
energy. In the final analysis all the energy

converted into heat and makes an independent contribution to the heating of the planet.

We know the mass of carbon burnt every year. Thus, it is easy to estimate the liberation of energy (the contribution of atomic power stations is so far rather moderate). The result is that the additional thermal power equals presently 2×10^{20} J per year or 6×10^{12} W. You see that today the power engineering of the world produces already a quantity of thermal power almost approaching by the order of magnitude the heat flux from the interior of the entire Earth (Ch. 4). Both these power fluxes are insignificant against the background of the power flux from the Sun. But what are we to expect in the future?

The global power consumption increases very rapidly. At the beginning of this century it was doubling every 50 years while presently it doubles every 20 years. The exhaustion of easily accessible oil and gas deposits will not halt the progress of power engineering. The coal will be available for a long time. In the distant future we expect to make use of the thermonuclear synthesis to generate power. Within the nearest fifty years oil-and-gas power engineering will be completely replaced by atomic power engineering. The store of uranium is quite ample. Already today we can use not only the isotope ^{235}U but also ^{238}U which is widely spread. Nuclear fuel is bred with the help of fast neutrons using the reaction $^{238}\mathrm{U} + \mathrm{n} \rightarrow {}^{239}\mathrm{Pu}$. Plutonium is a nuclear fuel in no way inferior to ^{235}U. Fast neutrons required for plutonium formation are produced in reactor in the course of decay of ^{235}U or ^{239}Pu. That is why this process is called nuclear fuel breeding.

In view of these prospects we may assume that the time during which the power engineering will double its production will remain in the future close to 20 years. This means that already in 200 years the power production on the Earth will reach 10^{16} W, i.e. 5% of the solar heat power received by the Earth. Such a power engineering will elevate the Earth's average temperature by a dangerous value of 3 degrees.

However, by that time our descendants will surely master an accurate control of the climate. From the technical point of view it can be done even today by way of changing the Earth's albedo and constructing dams in straits to control ocean currents. Realization of such projects has been so far impeded by our incompetence in forecasting expected effects. But the main thing is that we do not realize yet the necessity to join efforts of the entire humanity for the solution of such global problems.

The science about the Earth and the Earth's connection with space has always been deeply emotional. There is really no need to restrain enthusiasm originated by the feeling that we can understand the magnificent natural phenomena. Very recently Vladimir Vysotsky was singing:

Come, rain of stars inciting our souls
With ancient passion for the Earth and sea.